Accepted by Journal of Econometrics, Oct. 6, 2020

Diagnostic Tests for Homoskedasticity in Spatial Cross-Sectional or Panel Models

Badi H. Baltagi[a*], Alain Pirotte[b] and Zhenlin Yang[c]

[a] *Department of Economics and Center for Policy Research, Syracuse University, United States*

[b] *CRED, University Paris II Panthéon-Assas, France*

[c] *School of Economics, Singapore Management University, Singapore*

October 6, 2020

Abstract

We propose an *Adjusted Quasi-Score* (AQS) method for constructing tests for homoskedasticity in spatial econometric models. We first obtain an AQS function by adjusting the score-type function from the given model to achieve unbiasedness, and then develop an *Outer-Product-of-Martingale-Difference* (OPMD) estimate of its variance. In standard problems where a genuine (quasi) score vector is available, the AQS-OPMD method leads to finite sample improved tests over the usual methods. More importantly in non-standard problems where a genuine (quasi) score is not available and the usual methods fail, the proposed AQS-OPMD method provides feasible solutions. The AQS tests are formally derived and asymptotic properties examined for three representative models: spatial cross-sectional, static or dynamic panel models. Monte Carlo results show that the proposed AQS tests have good finite sample properties.

Keywords: Adjusted quasi-scores; Fixed effects; Heteroskedasticity; Incidental parameters; Martingale difference; Non-normality; Short dynamic panels; Spatial effects.

JEL Classifications: C12, C18, C21, C23.

*Corresponding author. E-mail: bbaltagi@maxwell.syr.edu.

1. Introduction

Spatial econometrics has experienced phenomenal growth in the last decade. Most of the estimation and inference methods are developed under the assumption that the errors in the model are homoskedastic. Therefore, it is important to test the validity of this assumption. Surprisingly, this type of test is largely unavailable in the spatial econometrics literature except Anselin (1988, p. 122), which is in stark contrast to the big literature on heteroskedasticity tests in the traditional (non-spatial) econometrics literature.

We in this paper endeavor to provide a general methodology for addressing the heteroskedasticity testing problem in spatial econometrics. We approach this problem by assuming, as in Breusch and Pagan (1979) and Anselin (1988), that the heteroskedasticity is induced by some exogenous variables such as certain covariates, the size of spatial units, and the number of neighbors, etc., through an unknown function so that the error variance has the form $\sigma^2 h(z'\alpha)$. The function h is such that $h(0) = 1$ and hence a test of heteroskedasticity becomes a test of null hypothesis $H_0 : \alpha = 0$. This leads naturally to the consideration of the score type of tests as their implementations require only the estimation of the simpler null models. Another advantage of such an approach is that by rejecting the null the 'source' of heteroskedasticity is 'identified' so that the model estimation and inference may proceed with heteroskedasticity of a *chosen form* of h in the spirit of Breusch and Pagan (1979).[1] However, the construction of classical LM tests depends upon the true score vector and information matrix, which are often unavailable for reasons given below, and hence the conventional methods fail. Moreover, the existence of spatial dependence often causes the LM tests to perform unsatisfactorily in finite samples even if the true scores and information matrix are available.[2] It is therefore highly desirable to have a general method that meets these challenges – able to provide not only the desired tests for homoskedasticity but also satisfactory ones.

This paper introduces a general method for constructing tests for homoskedasticity in spatial econometric models, namely the *Adjusted Quasi-Score* (AQS) method. We first obtain an AQS function by adjusting the score-type function from the given model to achieve unbiasedness/consistency, and then decompose the AQS function into a sum of vector mar-

[1]A strand of research that parallels our testing approach is the developments of estimation and inference procedures that are robust against cross-sectional heteroskedasticity of *unknown form*. See, among others, Lin and Lee (2010), Kelejian and Prucha (2010), Liu and Yang (2015b), Moscone and Tosetti (2011), Badinger and Egger (2011, 2015), Debarsy, Jin and Lee (2015), and Kuersteiner and Prucha (2020).

[2]See Baltagi and Yang (2013a,b), and Yang (2015a) for tests of spatial effects in linear or panel models.

tingale differences and hence an *Outer-Product-of-Martingale-Difference* (OPMD) estimate of its variance. In "standard problems", such as the spatial linear regression (SLR) model where a genuine (quasi) score vector is available, the AQS method leads to finite sample improved tests over the usual methods by adjusting the concentrated (quasi) score to remove the effect of estimating the linear and scale (*nuisance*) parameters. The role played by OPMD here is to provide a simple alternative in estimating the variance-covariance (VC) matrix of the AQS vector. However, the AQS-OPMD idea goes much beyond this – it provides feasible solutions to "non-standard problems" where the usual methods fail due to the lack of (i) a valid (quasi) score and (ii) a feasible method for VC matrix estimation.

For example, for a spatial panel data (SPD) model with fixed effects, the transformation method (Lee and Yu, 2010) cannot be used to remove the unobserved fixed effects if any of the following requirements is not met: balanced panel, additive fixed effects, time-invariant spatial weights, and time-invariant covariate and/or spatial effects. In these cases, the best we can have is the concentrated (quasi) likelihood/score function with the fixed effects (additive or interactive) being concentrated out. This concentrated (quasi) score vector does not lead to consistent estimation due to the well-known *incidental parameters problem* (IPP) of Neyman and Scott (1948). However, it can be adjusted to 'remove' the effect of estimating the *incidental parameters*, yielding an AQS vector that is unbiased and consistent.[3] Another (more important) example is the **dynamic** spatial panel data (DSPD) model with short panels. In this case, even if all the requirements as for SPD models are met, one is still unable to achieve either (i) or (ii) due to the well-known *initial values problem* (IVP), another form of IPP (see, e.g., Hsiao, 2014). In a dynamic panel data model (Hsiao et al., 2002) and a DSPD model with only spatial error (Su and Yang, 2015), initial values are modeled to give a 'full' likelihood function, but this approach cannot be applied to a general DSPD model as pointed out by Yang (2018a). In this case, the best we can have is the 'conditional' (quasi) likelihood/score treating the initial values as exogenously given. This conditional (quasi) score does not lead to consistent estimation unless the time dimension (T) is also large along with the cross-section dimension (n) but even in this case it incurs an asymptotic bias (Yu et al., 2008). However, it can again be adjusted, as in Yang (2018a) for a homoskedastic FE-DSPD model, to 'remove' the effect of IVP and to give a set of AQS functions for the heteroskedastic DSPD model that are unbiased and consistent – leading to valid AQS tests for homoskedasticity.

[3]This method finds root in Neyman and Scott (1948, Sec. 5) on *modified likelihood equations*.

Therefore, the proposed AQS-OPMD method does not only lead to homoskedasticity tests that improve the conventional LM tests, but also gives the desired tests (and their improved versions) for situations where IPP arises and hence the conventional methods fail. To provide a full illustration of the generality and versatility of the proposed method, the AQS tests are formally derived and asymptotic properties examined for three representative spatial models: cross-section, static and dynamic panels. Monte Carlo results show that the improved versions of the proposed tests have good finite sample properties.

The rest of the paper is organized as follows. Section 2 presents the score-type tests as well as their improved versions to test for homoskedasticity in a spatial cross-sectional model. Section 3 presents these tests as well as their improved versions for a static SPD model, and critically discusses their extensions to unbalanced panels, SPD models with time-varying spatial weights and regression/spatial coefficients, and SPD models with interactive fixed effects. Section 4 presents AQS tests and their improved versions for a DSPD model with fixed effects and short panels. Section 5 presents Monte Carlo results and Section 6 concludes. Proofs of all theorems are relegated to appendices which are available as supplemental material to save space. To facilitate practical applications of our tests, Matlab code is provided at the author's website: `http://cred.u-paris2.fr/PIROTTE/package.zip`.

2. Tests for Homoskedasticity: Spatial Linear Model

To help with a quick appreciation of the general ideas and principles behind the methodology adopted in constructing tests for homoskedasticity in spatial econometrics models, we start with a simple SLR (spatial linear regression) model to demonstrate how the simple OPMD-variant of the score test can be obtained, how it can be turned into a non-normality robust quasi score (QS) test, and how it can be adjusted to give finite sample improved score and QS tests. The SLR model takes the form:

$$Y_n = \lambda_1 W_{1n} Y_n + X_n \beta + U_n, \quad U_n = \lambda_2 W_{2n} U_n + V_n, \tag{2.1}$$

where Y_n is an $n \times 1$ vector of observations on the dependent variable, X_n is an $n \times p$ matrix of observations on the p explanatory variables. W_{jn}, $j = 1, 2$, are the $n \times n$ weighting matrices capturing the interactions among the n spatial units. They are assumed to be exogenously given with zero diagonal elements. $(\beta, \lambda_1, \lambda_2)$ are the common model parameters representing the covariate and spatial effects, respectively. V_n is an $n \times 1$ vector of independent disturbances which may exhibit unknown heteroskedasticity. In particular, the elements $\{v_{ni}\}$ of V_n have zero mean but heteroskedastic variances $\sigma^2 h(z'_{ni}\alpha)$ with the $k \times 1$ vectors z_{ni} and α being, respectively, the heteroskedasticity variables and the heteroskedasticity parameters.

The heteroskedasticity function $h(\cdot)$ is an unknown smooth function such that $h(0) = 1$. Thus, when $\alpha = 0$, the model becomes homoskedastic. A test for homoskedasticity against heteroskedasticity becomes as in Breusch and Pagan (1979) a test of:

$$H_0 : \alpha = 0 \;\text{ vs. } H_a : \alpha \neq 0. \tag{2.2}$$

The literature on heteroskedasticity testing is big, but largely for non-spatial models, except Anselin (1988, p. 122) who presents a test for a simple cross-sectional spatial error model. The variables in z_{ni} may contain some elements of the x_{ni}, the ith value of the set of regressors. In spatial models, z_{ni} may contain variables that relate to the spatial weight matrices, e.g., the number of non-zero elements in each row of W_{1n} (number of neighbors), *etc.* This makes the test of H_0 in the context of spatial models more appealing. In certain spatial models such as models with large group interaction (Lee, 2004, 2007), the elements of W_n depend on n and hence the values z_{ni} of the heteroskedasticity variables may also depend on n. The values of the exogenous variables x_{ni} are allowed to be n-dependent as well, because the models considered are allowed to contain spatial Durbin effects (Anselin, 1988, p. 40).[4]

2.1. Score and Quasi-Score Tests

Let $\mathcal{H}_n(\alpha) = \texttt{diag}(\{h(z'_{ni}\alpha)\})$, where $\texttt{diag}(\cdot)$ forms a diagonal matrix based on the given elements or a given vector. Denote $\theta = (\beta', \sigma^2, \lambda')'$, $\lambda = (\lambda_1, \lambda_2)'$ and $\psi = (\theta', \alpha')'$. The full Gaussian loglikelihood function for ψ is given by:

$$\begin{aligned} \ell_{\texttt{SLR}}(\psi) = \; & -\tfrac{n}{2}\log(2\pi\sigma^2) + \log|B_{1n}(\lambda_1)| + \log|B_{2n}(\lambda_2)| - \tfrac{1}{2}\log|\mathcal{H}_n(\alpha)| \\ & -\tfrac{1}{2\sigma^2} V'_n(\beta,\lambda)\mathcal{H}_n^{-1}(\alpha)V_n(\beta,\lambda), \end{aligned} \tag{2.3}$$

where $V_n(\beta,\lambda) = \mathbb{Y}_n(\lambda) - \mathbb{X}_n(\lambda_2)\beta$, $\mathbb{Y}_n(\lambda) = B_{2n}(\lambda_2)B_{1n}(\lambda_1)Y_n$, $\mathbb{X}_n(\lambda_2) = B_{2n}(\lambda_2)X_n$, and $B_{rn}(\lambda_r) = I_n - \lambda_r W_{rn}, r = 1, 2$. Maximizing $\ell_{\texttt{SLR}}(\psi)$ at the null, or $\ell_{\texttt{SLR}}(\psi)|_{H_0}$, gives the ML estimator (MLE) or quasi MLE (QMLE) $\tilde{\theta}_n$ of θ for the null model. Jin and Lee (2013) show that $\tilde{\theta}_n$ is $\sqrt{n}$-consistent and asymptotically normal under regularity conditions.

The (quasi) score function $S_{\texttt{SLR}}(\psi) = \frac{\partial}{\partial\psi}\ell_{\texttt{SLR}}(\psi)$ has the form:

[4] Model 2.1 can be extended by adding higher-order spatial lags in Y_n and U_n. See, e.g., Elhorst (2014) and Lee and Yu (2016) for discussions on spatial Durbin models and the associated issue of parameter identification.

$$
S_{\mathtt{SLR}}(\psi) = \begin{cases} \frac{1}{\sigma^2}\mathbb{X}_n'(\lambda_2)\mathcal{H}_n^{-1}(\alpha)V_n(\beta,\lambda), \\ \frac{1}{2\sigma^4}V_n'(\beta,\lambda)\mathcal{H}_n^{-1}(\alpha)V_n(\beta,\lambda) - \frac{n}{2\sigma^2}, \\ \frac{1}{\sigma^2}V_n'(\beta,\lambda)\mathcal{H}_n^{-1}(\alpha)B_{2n}(\lambda_2)W_{1n}Y_n - \mathrm{tr}[G_{1n}(\lambda_1)], \\ \frac{1}{\sigma^2}V_n'(\beta,\lambda)\mathcal{H}_n^{-1}(\alpha)G_{2n}(\lambda_2)V_n(\beta,\lambda) - \mathrm{tr}[G_{2n}(\lambda_2)], \\ \frac{1}{2\sigma^2}\dot{h}(z_{ni}'\alpha)\sum_{i=1}^n \Big[\big(\frac{v_{ni}^2(\beta,\lambda)}{h(z_{ni}'\alpha)} - \sigma^2\big)\frac{z_{ni}}{h(z_{ni}'\alpha)}\Big], \end{cases} \tag{2.4}
$$

where $G_{rn}(\lambda_r) = W_{rn}B_{rn}^{-1}(\lambda_r), r = 1, 2$, and $\dot{h}(x) = \frac{d}{dx}h(x)$. If the errors v_{ni} are normally distributed, then $\ell_{\mathtt{SLR}}(\psi)$ is the true loglikelihood and $S_{\mathtt{SLR}}(\psi)$ the true score, and the three classical principles, Wald, LR and Score (LM) can be used to test the general hypothesis, $H_0 : g(\psi) = 0$. In particular the score test takes the standard form $T_S = S_{\mathtt{SLR}}'(\tilde{\psi}_n)\Sigma_n^{-1}(\tilde{\psi}_n)S_{\mathtt{SLR}}(\tilde{\psi}_n)$, where $\tilde{\psi}_n$ is the null estimate of ψ, and $\Sigma_n(\psi_0) = -\mathrm{E}[\frac{\partial}{\partial\psi}S_{\mathtt{SLR}}(\psi_0)]$ is the information matrix. See Anselin (1988, Ch. 6) for a general discussion. For our homoskedasticity tests, $\tilde{\psi}_n = (\tilde{\theta}_n', 0_k')'$, where 0_k is a $k \times 1$ vector of zeros. Partition, according to (θ, α), $S_{\mathtt{SLR}}(\psi) = (S_{\mathtt{SLR},\theta}'(\psi), S_{\mathtt{SLR},\alpha}'(\psi))'$ and $\Sigma_n(\psi) = \{\Sigma_{n,\theta\theta}(\psi), \Sigma_{n,\theta\alpha}(\psi); \Sigma_{n,\alpha\theta}(\psi), \Sigma_{n,\alpha\alpha}(\psi)\}$ and note that $S_{\mathtt{SLR},\theta}(\tilde{\psi}_n) = 0$. The score tests reduces to $T_S = S_{\mathtt{SLR},\alpha}'(\tilde{\psi}_n)\Sigma_{n,\alpha\alpha.\theta}^{-1}(\tilde{\psi}_n)S_{\mathtt{SLR},\alpha}(\tilde{\psi}_n)$, where $\Sigma_{n,\alpha\alpha.\theta}(\psi) = \Sigma_{n,\alpha\alpha}(\psi) - \Sigma_{n,\alpha\theta}(\psi)\Sigma_{n,\theta\theta}^{-1}(\psi)\Sigma_{n,\theta\alpha}(\psi)$.[5] When the errors are non-normal, the score test may be invalid and its robust version is desired. Furthermore, the score tests may have poor finite sample performance, in particular in the presence of spatial dependence.

We first give an OPMD-variant of the score test in line with the general AQS-OPMD methodology adopted in this paper. The score at the null, $S_{\mathtt{SLR}}^\circ(\theta) = S_{\mathtt{SLR}}(\psi)|_{H_0}$, has a simpler form as $h(0) = 1$ and $\dot{h}(0)$ is a constant (see (B.1), Appendix B). Let θ_0 be the true value of θ. For ease of exposition, we drop the arguments of a quantity evaluated at the true parameter values, e.g., $V_n = V_n(\beta_0, \lambda_0)$, $B_{rn} = B_{rn}(\lambda_{r0})$, $G_{rn} = G_{rn}(\lambda_{r0})$, *etc.* The score vector $S_{\mathtt{SLR}}^\circ(\theta_0)$ is simplified into the following general form using $W_{1n}Y_n = G_{1n}B_{2n}^{-1}V_n + G_{1n}B_{2n}^{-1}\mathbb{X}_n\beta_0$:

$$
S_{\mathtt{SLR}}^\circ(\theta_0) = \begin{cases} \Pi_1'V_n, \\ V_n'\Phi_1V_n - \mathrm{E}(V_n'\Phi_1V_n), \\ V_n'\Phi_2V_n - \mathrm{E}(V_n'\Phi_2V_n) + V_n'\Pi_2, \\ V_n'\Phi_3V_n - \mathrm{E}(V_n'\Phi_3V_n), \\ \frac{1}{2\sigma_0^2}\dot{h}(0)\sum_{i=1}^n \big[(v_{ni}^2 - \sigma_0^2)z_{ni}\big], \end{cases} \tag{2.5}
$$

[5] See Anselin (1988, p. 122) for a special case of T_S where only the spatial error (λ_2) is present in the model. A popular variant of T_S is to use the observed information matrix, $-\frac{\partial}{\partial\psi}S_{\mathtt{SLR}}(\tilde{\psi}_n)$, in place of $\Sigma_n(\tilde{\psi}_n)$. Another one, the *Outer-Product-of-Gradients*, is not available as it requires the observations Y_{ni} to be independent.

where $\Pi_1 = \frac{1}{\sigma_0^2}\mathbb{X}_n$, and $\Pi_2 = \frac{1}{\sigma_0^2}B_{2n}G_{1n}B_{2n}^{-1}\mathbb{X}_n\beta_0$; $\Phi_1 = \frac{1}{2\sigma_0^4}I_n$, $\Phi_2 = \frac{1}{\sigma_0^2}B_{2n}G_{1n}B_{2n}^{-1}$, and $\Phi_3 = \frac{1}{\sigma_0^2}G_{2n}$; and 'E' corresponds to the null model and the true parameter θ_0.

For a general n-dimensional square matrix Φ_n, denote its strictly upper, strictly lower, and diagonal matrices by Φ_n^u, Φ_n^l and Φ_n^d. We have, $V_n'\Phi_n V_n = V_n'(\Phi_n^u + \Phi_n^l + \Phi_n^d)V_n = V_n'(\Phi_n^{u\prime} + \Phi_n^l + \Phi_n^d)V_n = V_n'\xi_n + V_n'\Phi_n^d V_n$, where $\xi_n = (\Phi_n^{u\prime} + \Phi_n^l)V_n$. It follows that,

$$V_n'\Phi_n V_n - \mathrm{E}(V_n'\Phi_n V_n) = \textstyle\sum_{i=1}^n \big[v_{ni}\xi_{ni} + (v_{ni}^2 - \sigma_0^2)\phi_{n,ii}\big] \equiv \sum_{i=1}^n g_{ni}(\theta_0), \tag{2.6}$$

where $\{\phi_{n,ii}\}$ are the diagonal elements of Φ_n. Note that the elements v_{ni} are independent and that ξ_{ni} is $\mathcal{F}_{n,i-1}$-measurable, where $\{\mathcal{F}_{n,i}\}$ is the increasing sequence of σ-fields generated by $(v_{n1}, \ldots, v_{ni})$, for $i = 1, \ldots, n$. It follows that $\{g_{ni}(\theta_0), \mathcal{F}_{n,i}\}$ form an MD sequence.

By (2.6), the quadratic forms in (2.5) have the MD decompositions: $V_n'\Phi_r V_n - \mathrm{E}(V_n'\Phi_r V_n) = \sum_{i=1}^n g_{r,ni}$, $r = 1, 2, 3$, where $g_{r,ni}$ is as $g_{ni}(\theta_0)$ in (2.6). These give an MD decomposition

$$S_{\mathtt{SLR}}^{\circ}(\theta_0) = \textstyle\sum_{i=1}^n \mathbf{g}_{ni}(\theta_0), \tag{2.7}$$

where $\mathbf{g}_{ni}(\theta_0) = (\mathbf{g}_{ni,\theta}'(\theta_0), \mathbf{g}_{ni,\alpha}'(\theta_0))'$, $\mathbf{g}_{ni,\theta}(\theta_0) = \{\Pi_{1i}'v_{ni},\ g_{1,ni},\ g_{2,ni} + \Pi_{2i}v_{ni},\ g_{3,ni}\}'$, and $\mathbf{g}_{ni,\alpha}(\theta_0) = \frac{1}{2\sigma_0^2}\dot{h}(0)(v_{ni}^2 - \sigma_0^2)z_{ni}$. Clearly, $\{\mathbf{g}_{ni}(\theta_0), \mathcal{F}_{n,i}\}_{i=1}^n$ form a vector MD sequence. Thus, $\Omega_n \equiv \mathrm{Var}[S_{\mathtt{SLR}}^{\circ}(\theta_0)] = \sum_{i=1}^n \mathrm{E}[\mathbf{g}_{ni}(\theta_0)\mathbf{g}_{ni}'(\theta_0)]$, and its sample analogue, $\sum_{i=1}^n \tilde{\mathbf{g}}_{ni}\tilde{\mathbf{g}}_{ni}'$, gives a consistent estimator in that $\frac{1}{n}\sum_{i=1}^n \tilde{\mathbf{g}}_{ni}\tilde{\mathbf{g}}_{ni}' - \frac{1}{n}\mathrm{Var}[S_{\mathtt{SLR}}^{\circ}(\theta_0)] = o_p(1)$, where $\tilde{\mathbf{g}}_{ni}$ is the null estimate of $\mathbf{g}_{ni}$.[6] This gives the OPMD-variant of the score test in two equivalent forms:

$$T_{\mathtt{SLR}} = \big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_{ni}'\big)\big(\sum_{i=1}^n \tilde{\mathbf{g}}_{ni}\tilde{\mathbf{g}}_{ni}'\big)^{-1}\big(\sum_{i=1}^n \tilde{\mathbf{g}}_{ni}\big), \tag{2.8}$$

$$= \big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\alpha}'\big)\big[\sum_{i=1}^n (\tilde{\mathbf{g}}_{ni,\alpha} - \tilde{K}_n\tilde{\mathbf{g}}_{ni,\theta})(\tilde{\mathbf{g}}_{ni,\alpha} - \tilde{K}_n\tilde{\mathbf{g}}_{ni,\theta})'\big]^{-1}\big(\sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\alpha}\big), \tag{2.9}$$

where $\tilde{K}_n = \big(\sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\alpha}\tilde{\mathbf{g}}_{ni,\theta}'\big)\big(\sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\theta}\tilde{\mathbf{g}}_{ni,\theta}'\big)^{-1}$.[7] Obviously, $T_{\mathtt{SLR}}$ is invariant to the unknown $\dot{h}(0)$ appearing in $\mathbf{g}_{ni,\alpha}(\theta_0)$, and hence it can be removed or simply set to 1.

The score test statistic, $T_{\mathtt{SLR}}$, is derived under the normality assumption and hence may not be robust. We now introduce a quasi score (QS) test allowing distributional misspecification. Note that the construction of the QS test depends upon $S_{\mathtt{SLR},\alpha}^{\circ}(\tilde{\theta}_n)$ and its VC matrix. Under mild conditions, Taylor expansions lead to the following **asymptotic MD representation**:

$$\tfrac{1}{\sqrt{n}}S_{\mathtt{SLR},\alpha}(\tilde{\theta}_n) = \tfrac{1}{\sqrt{n}}S_{\mathtt{SLR},\alpha}(\theta_0) - \tfrac{1}{\sqrt{n}}\Sigma_{n,\alpha\theta}\Sigma_{n,\theta\theta}^{-1}S_{\mathtt{SLR},\theta}(\theta_0) + o_p(1), \tag{2.10}$$

$$= \tfrac{1}{\sqrt{n}}\textstyle\sum_{i=1}^n (\mathbf{g}_{ni,\alpha} - \Gamma_n\mathbf{g}_{ni,\theta}) + o_p(1), \tag{2.11}$$

[6] Note that $\tilde{\mathbf{g}}_{ni}$ is obtained from $\mathbf{g}_{ni}(\theta_0)$ by replacing θ_0 by $\tilde{\theta}_n$ and $V_n = V_n(\beta_0, \lambda_0)$ by $\tilde{V}_n = V_n(\tilde{\beta}_n, \tilde{\lambda}_n)$.

[7] This first form is more general as it can be applied to test a general linear or nonlinear constraint on all parameters under normality. The latter form connects directly to the robust QS test given latter.

where $\Gamma_n = \Sigma_{n,\alpha\theta}\Sigma_{n,\theta\theta}^{-1}$. It is clear that $\{\mathbf{g}_{ni,\alpha} - \Gamma_n \mathbf{g}_{ni,\theta},\ \mathcal{F}_{n,i}\}$ form an MD sequence. Thus,

$$\mathrm{Var}[\tfrac{1}{\sqrt{n}} S_{\mathtt{SLR},\alpha}(\tilde{\theta}_n)] = \tfrac{1}{n}\textstyle\sum_{i=1}^n \mathrm{E}[(\mathbf{g}_{ni,\alpha} - \Gamma_n \mathbf{g}_{ni,\theta})(\mathbf{g}_{ni,\alpha} - \Gamma_n \mathbf{g}_{ni,\theta})'] + o(1), \tag{2.12}$$

leading immediately to an OPMD estimator of the $\mathrm{Var}[\frac{1}{\sqrt{n}} S_{\mathtt{SLR},\alpha}(\tilde{\theta}_n)]$, and an OPMD form of the QS test for homoskedasticity robust against non-normality:

$$T_{\mathtt{SLR}}^{\mathtt{r}} = \big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\alpha}'\big)\big[\textstyle\sum_{i=1}^n (\tilde{\mathbf{g}}_{ni,\alpha} - \tilde{\Gamma}_n \tilde{\mathbf{g}}_{ni,\theta})(\tilde{\mathbf{g}}_{ni,\alpha} - \tilde{\Gamma}_n \tilde{\mathbf{g}}_{ni,\theta})'\big]^{-1}\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\alpha}\big), \tag{2.13}$$

where $\tilde{\Gamma}_n = \tilde{\Sigma}_{n,\alpha\theta}\tilde{\Sigma}_{n,\theta\theta}^{-1}$, with $\tilde{\Sigma}_{n,\alpha\theta}$ and $\tilde{\Sigma}_{n,\theta\theta}$ being either the plug-in estimates of $\Sigma_{n,\alpha\theta}$ and $\Sigma_{n,\theta\theta}$, or simply $-\frac{\partial}{\partial\theta'} S_{\mathtt{SLR},\alpha}^{\circ}(\tilde{\theta}_n)$ and $-\frac{\partial}{\partial\theta'} S_{\mathtt{SLR},\theta}^{\circ}(\tilde{\theta}_n)$. The expressions for $\frac{\partial}{\partial\theta'} S_{\mathtt{SLR},\alpha}^{\circ}(\theta)$ and $\frac{\partial}{\partial\theta'} S_{\mathtt{SLR},\theta}^{\circ}(\theta)$ can easily be obtained from (2.4), and are given in Appendix B following the proof of Theorem 2.1. When the errors are normal, $\Sigma_n = \Omega_n$ (the information matrix equality, or IME). Then, $T_{\mathtt{SLR}}^{\mathtt{r}} = T_{\mathtt{SLR}}$ if the OPMD estimate of Ω_n is used. The asymptotic null distributions of the tests are established under the following standard regularity conditions:

Assumption 2.1. *The disturbances $\{v_{ni}, i = 1, \ldots, n\}$ are independent with means 0, variances $\sigma^2 h(z_{ni}'\alpha)$, and $E|v_{ni}|^{4+\epsilon} < \infty$ for some $\epsilon > 0$.*

Assumption 2.2. *The elements of X_n and $z_{ni}, i = 1, \ldots, n$, are nonstochastic and are uniformly bounded, and $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ exists and is nonsingular.*

Assumption 2.3. *W_{1n} and W_{2n} are uniformly bounded in both row and column sums in absolute value, and their diagonal elements are zero.*

Assumption 2.4. *B_{1n}^{-1} and B_{2n}^{-1} are uniformly bounded in both row and column sums in absolute value, uniformly in λ in a neighborhood of its true value.*

Theorem 2.1. *Under Assumptions 2.1-2.4, if $\tilde{\theta}_n$ is $\sqrt{n}$-consistent for θ_0 under H_0, and $\frac{1}{n}\Sigma_{n,\theta\theta}$ and $\frac{1}{n}\Omega_n$ are positive definite (p.d.) for large enough n, then, $T_{\mathtt{SLR}}^{\mathtt{r}}|_{H_0} \xrightarrow{D} \chi_k^2$ when the errors are either normal or non-normal; $T_{\mathtt{SLR}}|_{H_0} \xrightarrow{D} \chi_k^2$ when the errors are normal.*

The proof of Theorem 2.1 is given in Appendix B. The key tools used in the proof are the Central Limit Theorem (CLT) for Linear-Quadratic (LQ) forms of Kelejian and Prucha (2001), or its alternative version given in Lemma A.5, and the Weak Law of Large Numbers (WLLN) for MD arrays in, e.g., Davidson (1994, p. 299). The former leads to the result $\frac{1}{\sqrt{n}} S_{\mathtt{SLR}}^{\circ}(\theta_0) \xrightarrow{D} N(0, \lim_{n\to\infty} \frac{1}{n}\Omega_n)$, and the latter to the result $\frac{1}{n}\sum_{i=1}^n \tilde{\mathbf{g}}_{ni}\tilde{\mathbf{g}}_{ni}' - \frac{1}{n}\Omega_n \xrightarrow{p} 0$.

2.2. Adjusted Score and Adjusted Quasi-Score Tests

We have shown how the OPMD method leads to a simple variant of the score test, and how it leads quickly to a non-normality robust QS test. We now show how the AQS and OPMD methods together lead to finite sample improved tests. The β and σ^2 are linear and scale parameters, their constrained estimates given λ and α (nonlinear parameters) have analytical expressions, and hence they can be concentrated out to give the concentrated (quasi) score (CS or CQS) functions for λ and α.[8] However, the CS or CQS functions are no longer unbiased due to the additional variability inherited from the estimation of β and σ^2, which constitutes a major source of bias or size distortion in the subsequent estimation and testing of the nonlinear parameters (Yang, 2015b, Liu and Yang, 2015a). Therefore, an adjustment (bias-correction) of the CS or CQS functions would potentially lead to tests with improved finite sample performance (Baltagi and Yang, 2013a,b).

Working with (2.4), we obtain the constrained estimates of β and σ^2, given λ and $\alpha = 0$:

$$\tilde{\beta}_n(\lambda) = [\mathbb{X}_n'(\lambda_2)\mathbb{X}_n(\lambda_2)]^{-1}\mathbb{X}_n'(\lambda_2)\mathbb{Y}_n(\lambda) \text{ and } \tilde{\sigma}_n^2(\lambda) = \tfrac{1}{n}\mathbb{Y}_n'(\lambda)M_n(\lambda_2)\mathbb{Y}_n(\lambda), \quad (2.14)$$

where $M_n(\lambda_2) = I_n - \mathbb{X}_n(\lambda_2)[\mathbb{X}_n'(\lambda_2)\mathbb{X}_n(\lambda_2)]^{-1}\mathbb{X}_n'(\lambda_2)$. Substituting $\tilde{\beta}_n(\lambda)$ and $\tilde{\sigma}_n^2(\lambda)$ into the last three components of (2.4), we obtain the concentrated scores at the null:

$$S^c_{\text{SLR}}(\lambda,\alpha)|_{H_0} = \begin{cases} \tilde{\sigma}_n^{-2}(\lambda)\big\{\mathbb{Y}_n'(\lambda)M_n(\lambda_2)\big[B_{2n}(\lambda_2)G_{1n}(\lambda_1)B_{2n}^{-1}(\lambda_2) - \bar{G}_{1n}(\lambda_1)I_n\big]\mathbb{Y}_n(\lambda)\big\}, \\ \tilde{\sigma}_n^{-2}(\lambda)\big\{\mathbb{Y}_n'(\lambda)M_n(\lambda_2)\big[G_{2n}(\lambda_2) - \bar{G}_{2n}(\lambda_2)I_n\big]M_n(\lambda_2)\mathbb{Y}_n(\lambda)\big\}, \\ \tilde{\sigma}_n^{-2}(\lambda)\big\{\dot{h}(0)Z_n'\tilde{\zeta}_n(\lambda)\big\}, \end{cases} \quad (2.15)$$

where $Z_n = \{z_{ni}'\}_{n\times k}$, $\tilde{\zeta}_n(\lambda) = \frac{1}{2}\{v_{ni}^2(\tilde{\beta}_n(\lambda),\lambda) - \tilde{\sigma}_n^2(\lambda)\}$, and $\bar{G}_{rn}(\lambda_r) = \frac{1}{n}\text{tr}[G_{rn}(\lambda_r)]$, $r = 1, 2$.

Under mild conditions, the constrained QMLE $\tilde{\lambda}_n$ defined in Section 2.1 is equivalent to the solution of the following estimating equations: $\mathbb{Y}_n'(\lambda)M_n(\lambda_2)\big[B_{2n}(\lambda_2)G_{1n}(\lambda_1)B_{2n}^{-1}(\lambda_2) - \bar{G}_{1n}(\lambda_1)I_n\big]\mathbb{Y}_n(\lambda) = 0$ and $\mathbb{Y}_n'(\lambda)M_n(\lambda_2)\big[G_{2n}(\lambda_2) - \bar{G}_{2n}(\lambda_2)I_n\big]M_n(\lambda_2)\mathbb{Y}_n(\lambda) = 0$, obtained from the first two components of (2.15). However, neither estimation function has zero expectation, which constitutes a major source of finite sample bias of $\tilde{\lambda}_n$ (Yang, 2015b; Liu and Yang, 2015a), and a major source of size distortion for the tests of homoskedasticity (allowing spatial effects) constructed in Section 2.1, and the tests for spatial effects (Baltagi and Yang, 2013a,b; Yang, 2015a, 2018b). Noting that $\tilde{\sigma}_n^2(\lambda_0) \xrightarrow{p} \sigma_0^2$, we construct a test that potentially has better finite sample properties. This is done by working on the numerators of (2.15) or the quantities in the curling brackets, i.e., $\tilde{\sigma}_n^2(\lambda)S^c_{\text{SLR}}(\lambda,\alpha)|_{H_0}$.

Under H_0 and λ_0, we can easily see that $\mathbb{Y}_n'M_n(B_{2n}G_{1n}B_{2n}^{-1} - \bar{G}_{1n}I_n)\mathbb{Y}_n = V_n'\Phi_1V_n + \Pi'V_n$, and $\mathbb{Y}_n'M_n(G_{2n} - \bar{G}_{2n}I_n)M_n\mathbb{Y}_n = V_n'\Phi_2V_n$, where $\Pi = M_n(B_{2n}G_{1n}B_{2n}^{-1} - \bar{G}_{1n}I_n)\mathbb{X}_n\beta_0$, $\Phi_1 =$

[8]Nonlinear parameters are those whose estimates can only be obtained through numerical maximization or root-finding. Concentration simplifies the numerical process, especially when the dimension of β is large.

$M_n(B_{2n}G_{1n}B_{2n}^{-1} - \bar{G}_{1n}I_n)$ and $\Phi_2 = M_n(G_{2n} - \bar{G}_{2n}I_n)M_n$. These show that the expectations of the first two components of the numerator of (2.15) are, respectively, $\sigma_0^2 \text{tr}(\Phi_r), r = 1, 2$. Also, for the numerator of the last component of (2.15), we have $V_n(\tilde{\beta}_n(\lambda_0), \lambda_0) = M_n \mathbb{Y}_n = M_n V_n$. It follows that $\text{E}[v_{ni}^2(\tilde{\beta}_n(\lambda_0), \lambda_0)] = \text{E}[(M_{ni}V_n)^2] = \sigma_0^2 \sum_{j=1}^n M_{n,ij}^2 \equiv \sigma_0^2 m_i$, where M_{ni} denotes the ith row of M_n and $M_{n,ij}$ the ijth element of M_n. Define

$$\tilde{\zeta}_n^*(\lambda) = \tfrac{1}{2}\{\tfrac{1}{m_i(\lambda_2)} v_{ni}^2(\tilde{\beta}_n(\lambda), \lambda) - \tfrac{n}{n-p}\tilde{\sigma}_n^2(\lambda)\}_{n\times 1}. \tag{2.16}$$

The set of adjusted (concentrated) quasi scores (AQS) at H_0 thus have the simple form:

$$S_{\text{SLR}}^*(\lambda) = \begin{cases} \mathbb{Y}_n'(\lambda)\Phi_1(\lambda)\mathbb{Y}_n(\lambda) - \frac{n}{n-p}\tilde{\sigma}_n^2(\lambda)\text{tr}[\Phi_1(\lambda)], \\ \mathbb{Y}_n'(\lambda)\Phi_2(\lambda)\mathbb{Y}_n(\lambda) - \frac{n}{n-p}\tilde{\sigma}_n^2(\lambda)\text{tr}[\Phi_2(\lambda)], \\ Z_n'\tilde{\zeta}_n^*(\lambda). \end{cases} \tag{2.17}$$

It is easy to see that $\text{E}[S_{\text{SLR}}^*(\lambda_0)|_{H_0}] = 0$, and hence $S_{\text{SLR}}^*(\lambda_0)$ may lead to a potentially improved score-type test. To find its variance estimator, noting that $\tilde{\sigma}_n^2(\lambda_0) = \frac{1}{n}V_n'M_nV_n$, we have at H_0: $\mathbb{Y}_n'\Phi_1\mathbb{Y}_n - \frac{n}{n-p}\tilde{\sigma}_n^2(\lambda_0)\text{tr}(\Phi_1) = V_n'\Phi_1^*V_n + V_n'\Pi$, and $\mathbb{Y}_n'\Phi_2\mathbb{Y}_n - \frac{n}{n-p}\tilde{\sigma}_n^2(\lambda_0)\text{tr}(\Phi_2) = V_n'\Phi_2^*V_n$, where $\Phi_r^* = \Phi_r - \frac{1}{n-p}\text{tr}(\Phi_r)M_n, r = 1, 2$. Using (2.6) and noting that $\text{E}(V_n'\Phi_r^*V_n) = 0$, we have $V_n'\Phi_r^*V_n = \sum_{i=1}^n g_{r,ni}^*$, $r = 1, 2$, where $g_{r,ni}^* \equiv g_{r,ni}^*(\theta_0) = v_{ni}\xi_{ni}^* + (v_{ni}^2 - \sigma_0^2)\phi_{n,ii}^*$, $\{\xi_{ni}^*\} = \xi_n^* = (\Phi_r^{*u\prime} + \Phi_r^{*l})V_n$, and $\phi_{n,ii}^*$ are the diagonal elements of Φ_r^*. The $\{g_{r,ni}^*, \mathcal{F}_{n,i}\}$ form an MD sequence. The elements of $\tilde{\zeta}_n^*(\lambda)$ are asymptotically independent. Define,

$$\mathbf{g}_{ni}^*(\theta_0) = \{g_{1,ni}^* + \Pi_i v_{ni},\ g_{2,ni}^*,\ z_{ni}'\tilde{\zeta}_{ni}^*(\lambda_0)\}'. \tag{2.18}$$

Then, $S_{\text{SLR}}^*(\lambda_0) = \sum_{i=1}^n \mathbf{g}_{ni}^*(\theta_0)$, and it can be shown that

$$\tfrac{1}{n}\text{Var}[S_{\text{SLR}}^*(\lambda_0)] = \tfrac{1}{n}\sum_{i=1}^n \text{E}[\mathbf{g}_{ni}^*(\theta_0)\mathbf{g}_{ni}^{*\prime}(\theta_0)] + o(1).$$

A score-type test statistic, or the AQS test, for testing $H_0 : \alpha = 0$ takes the following form:

$$T_{\text{SLR}}^{\text{r}*} = \big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\alpha}^{*\prime}\big)\big[\textstyle\sum_{i=1}^n(\tilde{\mathbf{g}}_{ni,\alpha}^* - \tilde{\Gamma}_n^*\tilde{\mathbf{g}}_{ni,\lambda}^*)\big(\tilde{\mathbf{g}}_{ni,\alpha}^* - \tilde{\Gamma}_n^*\tilde{\mathbf{g}}_{ni,\lambda}^*\big)'\big]^{-1}\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\alpha}^*\big), \tag{2.19}$$

where $\tilde{\Gamma}_n^* = \tilde{\Sigma}_{n,\alpha\lambda}^*\tilde{\Sigma}_{n,\lambda\lambda}^{*-1}$, $\tilde{\Sigma}_{n,\alpha\lambda}^* = -\frac{\partial}{\partial\lambda}S_{\text{SLR},\alpha}^*(\tilde{\lambda}_n)$, and $\tilde{\Sigma}_{n,\lambda\lambda}^* = -\frac{\partial}{\partial\lambda}S_{\text{SLR},\lambda}^*(\tilde{\lambda}_n)$. These derivatives can be obtained from (2.17) after some tedious algebra, and their detailed expressions are given in Appendix B following the proof of Theorem 2.2.[9] When the errors are normally distributed, one could simply use $\tilde{\Sigma}_{n,\alpha\lambda}^* = \sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\alpha}^*\tilde{\mathbf{g}}_{ni,\lambda}^{*\prime}$ and $\tilde{\Sigma}_{n,\lambda\lambda}^* = \sum_{i=1}^n \tilde{\mathbf{g}}_{ni,\lambda}^*\tilde{\mathbf{g}}_{ni,\lambda}^{*\prime}$, leading to an adjusted score (AS) test, denoted by T_{SLR}^* for easy reference.[10]

[9] Numerical derivatives can be used in place of analytical ones: $\frac{\partial}{\partial\lambda_1}S_{\text{SLR}}^*(\lambda) = [S_{\text{SLR}}^*(\lambda + (\epsilon, 0)') - S_{\text{SLR}}^*(\lambda)]/\epsilon$ and $\frac{\partial}{\partial\lambda_2}S_{\text{SLR}}^*(\lambda) = [S_{\text{SLR}}^*(\lambda + (0, \epsilon)') - S_{\text{SLR}}^*(\lambda)]/\epsilon$, where ϵ is a small positive number, e.g., 0.00001.

[10] This is justified by an IME with respect to the underlining distribution (adjusted likelihood) that generates

The process of deriving $T^*_{\mathtt{SLR}}$ or $T^{\mathtt{r}*}_{\mathtt{SLR}}$ starts from the concentrated score where the variability from the estimation of β and σ^2 is captured. Then one recenters the numerator of the concentrated scores, and then rescales the 'recentered' score. Thus, these tests are expected to perform better in finite samples than $T_{\mathtt{SLR}}$ or $T^{\mathtt{r}}_{\mathtt{SLR}}$. Note that unlike the case with joint scores, $S^*_{\mathtt{SLR},\lambda}(\tilde{\lambda}_n)$ is not identically zero, as $\tilde{\lambda}_n$ is not the solution of the estimating equation $S^*_{\mathtt{SLR},\lambda}(\lambda) = 0$. In this case, an adjusted estimator that solves the AQS equations, i.e.,

$$\tilde{\lambda}^*_n = \arg\{S^*_{\mathtt{SLR},\lambda}(\lambda) = 0\}, \tag{2.20}$$

should be used to ensure good finite sample performance of the AQS test. This is confirmed by the Monte Carlo results presented in Section 5. The asymptotic null behavior of $T^*_{\mathtt{SLR}}$ and $T^{\mathtt{r}*}_{\mathtt{SLR}}$ is summarized in the following theorem:

Theorem 2.2. *Under the assumptions of Theorem 2.1,* $T^{\mathtt{r}*}_{\mathtt{SLR}}|_{H_0} \xrightarrow{D} \chi^2_k$ *when the errors are either normal or non-normal; and* $T^*_{\mathtt{SLR}}|_{H_0} \xrightarrow{D} \chi^2_k$ *when the errors are normal.*

3. Tests for Homoskedasticity: Spatial Panel Data Model

As indicated in the introduction, tests for homoskedasticity for spatial panel data (SPD) models are largely unavailable, even the three classical tests under normality. In this section, we first demonstrate how the AQS-OPMD methodology quickly leads to the desired tests using a 'standard' SPD model with fixed effects (FE), and then (more importantly) we demonstrate how it handles the complications caused by the existence of *incidental parameters*, the FE, when panels become 'non-standard' and hence the conventional method cannot be applied.

3.1. Homoskedasticity Tests: Standard FE-SPD Model

The 'standard' FE-SPD model, with a balanced panel, time-invariant spatial weights, time-invariant parameters and additive fixed effects, takes the form:

$$Y_{nt} = \lambda_1 W_{1n} Y_{nt} + X_{nt}\beta + \mu_n + a_t 1_n + U_{nt}, \;\; U_{nt} = \lambda_2 W_{2n} U_{nt} + V_{nt}, \;\; t = 1, 2, \ldots, T, \tag{3.1}$$

where, in period t, Y_{nt} denotes the $n \times 1$ vector of observations on the dependent variable, X_{nt} the $n \times p$ matrix of observations on p nonstochastic, unit and time varying regressors, and V_{nt} the $n \times 1$ vector of idiosyncratic errors $\{v_{it}\}$, which are independent across i and t

the AQS (2.17). Alternatively, the generalized IME can be applied to give $\tilde{\Sigma}^*_{n,\alpha\lambda} = \sum_{i=1}^n \tilde{\mathbf{g}}^*_{ni,\alpha}\tilde{\mathbf{g}}'_{ni,\lambda}$ and $\tilde{\Sigma}^*_{n,\lambda\lambda} = \sum_{i=1}^n \tilde{\mathbf{g}}^*_{ni,\lambda}\tilde{\mathbf{g}}'_{ni,\lambda}$, where $\tilde{\mathbf{g}}_{ni,\lambda}$ is the restricted estimate of the λ-element of the full $\mathbf{g}_{ni}$ in (2.7). However, the numerical results show that the former performs better in finite samples.

with means 0 and variances $\sigma^2 h(z_{ni}'\alpha)$. The parameters β, λ_1 and λ_2 are defined in the same way as in model (2.1), and μ_n represents the vector of unit-specific effects and $\{a_t\}$ the time-specific effects, which may correlate arbitrarily with unit and time varying regressors – the *fixed effects*. The two appear in the model additively giving rise to the additive FE model.[11]

Again, a test for homoskedasticity across the cross-section dimension corresponds to the test of the null hypothesis $H_0 : \alpha = 0$. For ease of exposition, we focus on the SDP model unit-specific FE only, i.e., dropping the time-specific FE.[12] We present an OPMD-variant of the score test, and a non-normality robust quasi score (QS) test. Then, we give a pair of adjusted score (AS) and adjusted quasi score (AQS) tests with finite sample improvements. Formal asymptotic theories are presented with the proofs relegated to Appendix C.

3.1.1. The ML or QML estimation

The ML or QML estimation of the FE-SPD model under $H_0 : \alpha = 0$ proceeds with the transformation approach followed by Lee and Yu (2010) and Yang et al. (2016). To eliminate the individual effects, define $J_T = (I_T - \frac{1}{T} l_T l_T')$ and let $[F_{T,T-1}, \frac{1}{\sqrt{T}} l_T]$ be the orthonormal eigenvector matrix of J_T, where $F_{T,T-1}$ is the $T \times (T-1)$ submatrix corresponding to the eigenvalues of one, I_T is a $T \times T$ identity matrix and l_T is a $T \times 1$ vector of ones. For any $n \times T$ matrix $[A_{n1}, \ldots, A_{nT}]$, define the $n \times (T-1)$ transformed matrix as $[A_{n1}^*, \ldots, A_{n,T-1}^*] = [A_{n1}, \ldots, A_{nT}]F_{T,T-1}$. This leads to the transformed vectors: Y_{nt}^*, U_{nt}^*, V_{nt}^*, and $X_{nt,j}^*$ for the jth regressor, for $t = 1, \ldots, T-1$. Let $X_{nt}^* = [X_{nt,1}^*, X_{nt,2}^*, \ldots, X_{nt,k}^*]$. We have:

$$Y_{nt}^* = \lambda_1 W_{1n} Y_{nt}^* + X_{nt}^* \beta + U_{nt}^*,\ U_{nt}^* = \lambda_2 W_{2n} U_{nt}^* + V_{nt}^*,\ t = 1, \ldots, T-1. \tag{3.2}$$

After the transformation, the effective sample size becomes $N = n(T-1)$. Letting $\mathbf{Y}_N = (Y_{n1}^{*\prime}, \ldots, Y_{n,T-1}^{*\prime})'$, and similarly for $\mathbf{U}_N$, $\mathbf{V}_N$ and $\mathbf{X}_N$. Denoting $\mathbf{W}_{rN} = I_{T-1} \otimes W_{rn}, r = 1, 2$, where $\otimes$ is the Kronecker product, the transformed model (3.2) is compactly written as:

$$\mathbf{Y}_N = \lambda_1 \mathbf{W}_{1N} \mathbf{Y}_N + \mathbf{X}_N \beta + \mathbf{U}_N,\ \ \mathbf{U}_N = \lambda_2 \mathbf{W}_{2N} \mathbf{U}_N + \mathbf{V}_N, \tag{3.3}$$

which is identical in form to the SLR model. Hence, the estimation of the FE-SPD model is similar. The key difference is that the elements $\{v_{it}^*\}$ of the transformed error vector $\mathbf{V}_N$ may not be totally independent unless the original errors are independent and normal. When the

[11]Similar to the SLR model, the FE-SPD model can also be extended by adding spatial Durbin terms, higher-order spatial lags of response and disturbance.

[12]At the end of this subsection, we outline how the results can be extended to allow for time-specific FE. We further discuss how the results can be extended to allow for time-wise heteroskedasticity.

original errors are independent but non-normal, $\{v^*_{it}\}$ are independent across i by definition but only uncorrelated across t, as $(V^{*\prime}_{n1}, \ldots, V^{*\prime}_{n,T-1})' = (F'_{T,T-1} \otimes I_n)(V'_{n1}, \ldots, V'_{nT})'$, and

$$\begin{aligned} & E(V^{*\prime}_{n1}, \ldots, V^{*\prime}_{n,T-1})'(V^{*\prime}_{n1}, \ldots, V^{*\prime}_{n,T-1}) \\ = \; & \sigma^2(F'_{T,T-1} \otimes I_n)(I_T \otimes \mathcal{H}_n(\alpha))(F_{T,T-1} \otimes I_n) \\ = \; & \sigma^2(I_{T-1} \otimes \mathcal{H}_n(\alpha)) \equiv \sigma^2 \mathbf{H}_N(\alpha), \end{aligned} \tag{3.4}$$

where $\mathcal{H}_n(\alpha)$ is defined in Section 2. It follows that the full quasi Gaussian loglikelihood function for $\psi = (\beta', \sigma^2, \lambda', \alpha')'$ (required for the derivation of the score-type tests later) is,

$$\begin{aligned} \ell_{\mathtt{SPD}}(\psi) = \; & -\tfrac{N}{2}\log(2\pi\sigma^2) + \log|\mathbf{B}_{1N}(\lambda_1)| + \log|\mathbf{B}_{2N}(\lambda_2)| \\ & -\tfrac{1}{2}\log|\mathbf{H}_N(\alpha)| - \tfrac{1}{2\sigma^2}\mathbf{V}'_N(\beta,\lambda)\mathbf{H}^{-1}_N(\alpha)\mathbf{V}_N(\beta,\lambda), \end{aligned} \tag{3.5}$$

where $\mathbf{V}_N(\beta,\lambda) = \mathbb{Y}_N(\lambda) - \mathbb{X}_N(\lambda_2)\beta$, $\mathbb{Y}_N(\lambda) = \mathbf{B}_{2N}(\lambda_2)\mathbf{B}_{1N}(\lambda_1)\mathbf{Y}_N$, $\mathbb{X}_N(\lambda_2) = \mathbf{B}_{2N}(\lambda_2)\mathbf{X}_N$, and $\mathbf{B}_{rN}(\lambda_1) = I_N - \lambda_r \mathbf{W}_{rN}, r = 1, 2$. Maximizing $\ell_{\mathtt{SPD}}(\psi)|_{H_0}$ gives the (Q)MLE $\tilde{\theta}_N$ of $\theta = (\beta', \sigma^2, \lambda')'$ in the null model, which is $\sqrt{N}$-consistent as shown in Lee and Yu (2010).

3.1.2. The score and quasi-score tests

The same idea as in Section 2.1 can be followed to give a score or QS test of homoskedasticity in the FE-SPD model. However, it should be noted that when the original errors are non-normal, the transformed errors are independent along the cross-sectional dimension only, not along the time dimension although they are still uncorrelated. While this makes the derivations of the results and the proof of the theorems more difficult, it emphasizes the advantage of the proposed OPMD method. This is because under the transformed QML approach, the explicit VC matrix of the score vector involves the unknown 3rd and 4th moments of the original errors v_{it}, but only the estimated residuals on the transformed scale are available.

The (quasi) score function $S_{\mathtt{SPD}}(\psi) = \frac{\partial}{\partial \psi}\ell_{\mathtt{SPD}}(\psi)$ can be easily derived, which gives $S^{\circ}_{\mathtt{SPD}}(\theta) = S_{\mathtt{SPD}}(\psi)|_{H_0}$ as follows, using that facts that $h(0) = 1$ and $\dot{h}(0)$ is a constant free of i and t:

$$S^{\circ}_{\mathtt{SPD}}(\theta) = \begin{cases} \frac{1}{\sigma^2}\mathbb{X}'_N(\lambda_2)\mathbf{V}_N(\beta,\lambda), \\ \frac{1}{2\sigma^4}\mathbf{V}'_N(\beta,\lambda)\mathbf{V}_N(\beta,\lambda) - \frac{N}{2\sigma^2}, \\ \frac{1}{\sigma^2}\mathbf{V}'_N(\beta,\lambda)\mathbf{B}_{2N}(\lambda_2)\mathbf{W}_{1N}\mathbf{Y}_N - \mathrm{tr}[\mathbf{G}_{1N}(\lambda_1)], \\ \frac{1}{\sigma^2}\mathbf{V}'_N(\beta,\lambda)\mathbf{G}_{2N}(\lambda_2)\mathbf{V}_N(\beta,\lambda) - \mathrm{tr}[\mathbf{G}_{2N}(\lambda_2)], \\ \frac{1}{2\sigma^2}\dot{h}(0)\sum_{t=1}^{T-1}\sum_{i=1}^{n}\left[v^{*2}_{it}(\beta,\lambda) - \sigma^2\right]z_{ni}. \end{cases} \tag{3.6}$$

where $\mathbf{G}_{rN}(\lambda_r) = I_{T-1} \otimes G_{rn}(\lambda_r)$, $r = 1, 2$. At the null and the true parameter values

$\theta_0 = (\beta_0', \sigma_0^2, \lambda_0')'$, $\mathbf{H}_N^{-1}(0) = I_N$, $\mathbf{V}_N(\beta_0, \lambda_0) = \mathbf{V}_N$, $\mathbf{B}_{2N}(\lambda_{20}) = \mathbf{B}_{2N}$, and $\mathbf{G}_{rN}(\lambda_{r0}) = \mathbf{G}_{rN}$. To derive a variance estimator, as in (2.5) we express $S^{\circ}_{\mathtt{SPD}}(\theta_0)$ in terms of $\mathbf{V}_N$ and θ_0:

$$S^{\circ}_{\mathtt{SPD}}(\theta_0) = \begin{cases} \Pi_1' \mathbf{V}_N, \\ \mathbf{V}_N' \mathbf{\Phi}_1 \mathbf{V}_N - \mathrm{E}(\mathbf{V}_N' \mathbf{\Phi}_1 \mathbf{V}_N), \\ \mathbf{V}_N' \mathbf{\Phi}_2 \mathbf{V}_N - \mathrm{E}(\mathbf{V}_N' \mathbf{\Phi}_2 \mathbf{V}_N) + \mathbf{V}_N' \Pi_2, \\ \mathbf{V}_N' \mathbf{\Phi}_3 \mathbf{V}_N - \mathrm{E}(\mathbf{V}_N' \mathbf{\Phi}_3 \mathbf{V}_N), \\ \frac{1}{2\sigma_0^2} \dot{h}(0) \sum_{t=1}^{T-1} \sum_{i=1}^{n} (v_{it}^{*2} - \sigma_0^2) z_{ni}, \end{cases} \tag{3.7}$$

where $\Pi_1 = \frac{1}{\sigma_0^2} \mathbb{X}_N(\lambda_2)$, $\Pi_2 = \frac{1}{\sigma_0^2} \mathbf{B}_{2N} \mathbf{G}_{1N} \mathbf{B}_{2N}^{-1} \mathbb{X}_N(\lambda_2)\beta$, $\mathbf{\Phi}_1 = \frac{1}{2\sigma_0^4} I_N$, $\mathbf{\Phi}_2 = \frac{1}{\sigma_0^2} \mathbf{B}_{2N} \mathbf{G}_{1N} \mathbf{B}_{2N}^{-1}$, and $\mathbf{\Phi}_3 = \frac{1}{\sigma_0^2} \mathbf{G}_{2N}$. In an identical way leading to (2.7), we can write $S^{\circ}_{\mathtt{SPD}}(\theta_0) = \sum_{j=1}^{N} \mathbf{g}_{Nj}(\theta_0)$, where j $(= 1, \ldots, N)$ is the combined index for (i, t) with $i = 1, \ldots, n$ for each $t = 1, \ldots, T-1$, and the detailed expression of $\mathbf{g}_{Nj}(\theta_0)$ is given in (C.1) of Appendix C.

If the original errors $\{v_{it}\}$ are *iid* normal, then the transformed errors $\{v_{it}^*\}$ or $\{v_j^*\}$ are *iid* normal, and based on the same reasoning as for the SLR model, $\{\mathbf{g}_{Nj}(\theta)\}$ form an MD sequence with respect to the increasing σ-fields $\{\mathcal{F}_{N,j}\}$ generated by $(v_1^*, \ldots, v_j^*)$. Thus, a consistent estimator for $\mathbf{\Omega}_N = \frac{1}{N} \mathrm{Var}[S^{\circ}_{\mathtt{SPD}}(\theta_0)]$ is

$$\widetilde{\mathbf{\Omega}}_N = \tfrac{1}{N} \textstyle\sum_{j=1}^{N} \tilde{\mathbf{g}}_{Nj} \tilde{\mathbf{g}}_{Nj}', \tag{3.8}$$

where $\tilde{\mathbf{g}}_{Nj} = \mathbf{g}_{Nj}(\tilde{\theta}_N)$. The OPMD-version of the score statistic, for testing H_0: $\alpha = 0$, has two equivalent forms identical to those in (2.8) and (2.9) for the SLR model:

$$T_{\mathtt{SPD}} = \big(\textstyle\sum_{j=1}^{N} \tilde{\mathbf{g}}_{Nj}' \big) \big[\sum_{j=1}^{N} \tilde{\mathbf{g}}_{Nj} \tilde{\mathbf{g}}_{Nj}' \big]^{-1} \big(\sum_{j=1}^{n} \tilde{\mathbf{g}}_{Nj} \big), \tag{3.9}$$

$$= \big(\textstyle\sum_{j=1}^{N} \tilde{\mathbf{g}}_{Nj,\alpha}' \big) \big[\sum_{j=1}^{N} (\tilde{\mathbf{g}}_{Nj,\alpha} - \tilde{K}_N \tilde{\mathbf{g}}_{Nj,\theta}) (\tilde{\mathbf{g}}_{Nj,\alpha} - \tilde{K}_N \tilde{\mathbf{g}}_{Nj,\theta})' \big]^{-1} \big(\sum_{j=1}^{n} \tilde{\mathbf{g}}_{Nj,\alpha} \big), \tag{3.10}$$

where $\tilde{K}_N = \big(\sum_{J=1}^{n} \tilde{\mathbf{g}}_{Nj,\alpha} \tilde{\mathbf{g}}_{Nj,\theta}' \big) \big(\sum_{j=1}^{n} \tilde{\mathbf{g}}_{Nj,\theta} \tilde{\mathbf{g}}_{Nj,\theta}' \big)^{-1}$. Again, the unknown constant $\dot{h}(0)$ appearing in the score element for α cancels out, and hence it can simply be set to 1.

If $\{v_{it}\}$ are *iid* but not normal, $\{v_{it}^*\}$ or $\{v_j^*\}$ are not guaranteed to be totally independent in the sense that there may exist a higher-order dependence among $\{v_{it}^*\}$. If this higher-order dependence does not affect the asymptotic properties of the OPMD estimate given in (3.8), then, similar to the QS test given in (2.13), a QS test for homoskedasticity in the FE-SPD model, allowing the errors to be non-normally distributed, can be obtained by replacing $\tilde{K}_N$ in (3.10) by $\widetilde{\mathbf{\Gamma}}_N = \tilde{\mathbf{\Sigma}}_{N,\alpha\theta} \tilde{\mathbf{\Sigma}}_{N,\theta\theta}^{-1}$, where $\mathbf{\Sigma}_{N,\alpha\theta} = -\mathrm{E}[\frac{\partial}{\partial \theta'} S^{\circ}_{\mathtt{SPD},\alpha}(\theta_0)]$ and $\mathbf{\Sigma}_{N,\theta\theta} = -\mathrm{E}[\frac{\partial}{\partial \theta'} S^{\circ}_{\mathtt{SPD},\theta}(\theta_0)]$. The **resulting test** is denoted by $T^{\mathtt{r}}_{\mathtt{SPD}}$ for easy reference. The analytical expressions for

$\frac{\partial}{\partial\theta'} S^{\circ}_{\mathtt{SPD},\alpha}(\theta)$ and $\frac{\partial}{\partial\theta'} S^{\circ}_{\mathtt{SPD},\theta}(\theta)$ are given in Appendix C.

However, we show in Appendix C that the correlation between v^*_{it} and v^{*2}_{is} and in particular the correlation between v^{*2}_{it} and v^{*2}_{is} induce correlation between $\mathbf{g}_{N,it}(\theta_0)$ and $\mathbf{g}_{N,is}(\theta_0)$, $t \neq s$, which may not be ignored when the skewness and excess kurtosis of v_{it} are not zero. It is shown in Appendix C that $\mathbf{\Omega}_N = \sum_{j=1}^N \mathrm{E}(\mathbf{g}_{Nj}\mathbf{g}'_{Nj}) + \sum_{j=1}^N \sum_{k(\neq j)=1}^N \mathrm{E}(\mathbf{g}_{Nj}\mathbf{g}'_{Nk}) = \sum_{j=1}^N \mathrm{E}(\mathbf{g}_{Nj}\mathbf{g}'_{Nj} + \mathbf{d}_{Nj}\mathbf{d}^{\circ\prime}_{Nj})$, and hence an extended OPMD estimate of $\mathbf{\Omega}_N$, taking into account the possible correlation between $\mathbf{g}_{N,it}(\theta_0)$ and $\mathbf{g}_{N,is}(\theta_0)$, $t \neq s$, is given as follows:

$$\widetilde{\mathbf{\Omega}}^{\mathrm{r}}_N = \textstyle\sum_{j=1}^N \big(\tilde{\mathbf{g}}_{Nj}\tilde{\mathbf{g}}'_{Nj} + \tilde{\mathbf{d}}_{Nj}\tilde{\mathbf{d}}^{\circ\prime}_{Nj}\big), \tag{3.11}$$

where $\tilde{\mathbf{g}}_{Nj}$ is given in (3.8), and $\tilde{\mathbf{d}}_{Nj}$ and $\tilde{\mathbf{d}}^{\circ}_{Nj}$ are the null estimates of $\mathbf{d}_{Nj}$ and $\mathbf{d}^{\circ}_{Nj}$, with $\mathbf{d}_{N,it} = \big\{\Pi'_{1,it}v^*_{it}, (v^{*2}_{it} - \sigma_0^2)\phi_{1,it}, (v^{*2}_{it} - \sigma_0^2)\phi_{2,it} + \Pi_{2,it}v^*_{it}, (v^{*2}_{it} - \sigma_0^2)\phi_{3,it}, \frac{1}{2\sigma_0^2} z'_{ni}(v^{*2}_{it} - \sigma_0^2)\big\}'$, and $\mathbf{d}^{\circ}_{N,it} = \sum_{s(\neq t)=1}^{T-1} \mathbf{d}_{N,is}$. The coefficients $\phi_{r,it}$ represent the diagonal elements of $\mathbf{\Phi}_r$, $r = 1, 2, 3$.

Now, similar to (2.10) an asymptotic expansion for $S^{\circ}_{\mathtt{SPD},\alpha}(\tilde\theta_N)$ leads to

$$\mathrm{Var}[S^{\circ}_{\mathtt{SPD},\alpha}(\tilde\theta_N)] = \mathbf{\Omega}_{N,\alpha\alpha} - \mathbf{\Gamma}_N\mathbf{\Omega}_{N,\theta\alpha} - \mathbf{\Omega}_{N,\alpha\theta}\mathbf{\Gamma}'_N + \mathbf{\Gamma}_N\mathbf{\Omega}_{N,\theta\theta}\mathbf{\Gamma}'_N + o(N), \tag{3.12}$$

and using the expression for $\mathbf{\Omega}_N$ given above, it can be expressed as

$$\mathrm{Var}[S^{\circ}_{\mathtt{SPD},\alpha}(\tilde\theta_N)] = \textstyle\sum_{j=1}^N \mathrm{E}(\mathbf{s}_{Nj}\mathbf{s}'_{Nj} + \mathbf{f}_{Nj}\mathbf{f}^{\circ\prime}_{Nj}) + o(N), \tag{3.13}$$

where $\mathbf{s}_{Nj,\alpha} = \mathbf{g}_{Nj,\alpha} - \mathbf{\Gamma}_N\mathbf{g}_{Nj,\theta}$, $\mathbf{f}_{Nj,\alpha} = \mathbf{d}_{Nj,\alpha} - \mathbf{\Gamma}_N\mathbf{d}_{Nj,\theta}$, and $\mathbf{f}^{\circ}_{Nj,\alpha} = \mathbf{d}^{\circ}_{Nj,\alpha} - \mathbf{\Gamma}_N\mathbf{d}^{\circ}_{Nj,\theta}$; and $(\mathbf{g}'_{Nj,\theta}, \mathbf{g}'_{Nj,\alpha})' = \mathbf{g}_{Nj}$, $(\mathbf{d}'_{Nj,\theta}, \mathbf{d}'_{Nj,\alpha})' = \mathbf{d}_{Nj}$, and $(\mathbf{d}^{\circ\prime}_{Nj,\theta}, \mathbf{d}^{\circ\prime}_{Nj,\alpha})' = \mathbf{d}^{\circ}_{Nj}$. Either of these two forms can be used to construct the test, and we choose the latter to be in line with the OPMD notion. A test statistic fully robust against non-normality is thus:

$$T^{\mathtt{rr}}_{\mathtt{SPD}} = \big(\textstyle\sum_{j=1}^N \tilde{\mathbf{g}}'_{Nj,\alpha}\big)\big[\sum_{j=1}^N (\tilde{\mathbf{s}}_{Nj}\tilde{\mathbf{s}}'_{Nj} + \tilde{\mathbf{f}}_{Nj}\tilde{\mathbf{f}}^{\circ\prime}_{Nj})\big]^{-1}\big(\sum_{j=1}^N \tilde{\mathbf{g}}_{Nj,\alpha}\big), \tag{3.14}$$

where $\tilde{\mathbf{s}}_{Nj}$, $\tilde{\mathbf{f}}_{Nj}$, $\tilde{\mathbf{f}}^{\circ}_{Nj}$ are estimates of $\mathbf{s}_{Nj}$, $\mathbf{f}_{Nj}$ and $\mathbf{f}^{\circ}_{Nj}$ at the null.

In a similar manner as for the SLR model, asymptotic normality of $S^{\circ}_{\mathtt{SPD}}(\theta_0)$ can be established using the CLT for LQ forms given in Lemma A.5, and the consistency of the variance estimator can be established using the WLLN for MD arrays in Davidson (1994, p. 229).

Theorem 3.1. *Extending Assumption 2.1 to $\{v_{it}\}$ and Assumption 2.2 to $\mathbf{X}_N$, and keeping Assumptions 2.3 and 2.4, if $\tilde\theta_N$ is $\sqrt{N}$-consistent under H_0, and $\frac{1}{N}\mathbf{\Sigma}_{N,\theta\theta}$ and $\frac{1}{N}\mathbf{\Omega}_N$ are p.d. for large enough N, then as n goes large (where T can be large or fixed),*

(i) $T_{\mathtt{SPD}}|_{H_0} \xrightarrow{D} \chi^2_k$ when the errors are normal;

(ii) $T^{\mathtt{rr}}_{\mathtt{SPD}}|_{H_0} \xrightarrow{D} \chi^2_k$ when the errors are either normal or non-normal.

3.1.3. The adjusted score and adjusted quasi-score tests

Following the same idea of Section 2.2, one can derive a potentially improved test for homoskedasticity for the FE-SPD model by working with the concentrated (quasi) scores and treating the elements of $\mathbf{V}_N$ as completely independent (recall: they are independent across i but only uncorrelated across t in general unless the original errors are normal). Referring to Sections 3.1.1 and 3.1.2 for notation, the constrained estimates of β and σ^2 are

$$\tilde{\beta}_N(\lambda) = [\mathbb{X}_N'(\lambda_2)\mathbb{X}_N(\lambda_2)]^{-1}\mathbb{X}_N'(\lambda_2)\mathbb{Y}_N(\lambda) \text{ and } \tilde{\sigma}_N^2(\lambda) = \tfrac{1}{N}\mathbb{Y}_N'(\lambda)\mathbf{M}_N(\lambda_2)\mathbb{Y}_N(\lambda),$$

where $\mathbf{M}_N(\lambda_2) = I_N - \mathbb{X}_N(\lambda_2)[\mathbb{X}_N'(\lambda_2)\mathbb{X}_N(\lambda_2)]^{-1}\mathbb{X}_N'(\lambda_2)$. Similar to (2.16), define

$$\tilde{\boldsymbol{\zeta}}_N^*(\lambda) = \tfrac{1}{2}\big\{\tfrac{1}{\mathbf{m}_j(\lambda_2)}v_{Nj}^2(\tilde{\beta}_N(\lambda),\lambda) - \tfrac{N}{N-p}\tilde{\sigma}_N^2(\lambda),\ j=1,\cdots,N\big\}_{N\times 1}, \tag{3.15}$$

where $\mathbf{m}_j(\lambda_2) = \sum_{\ell=1}^N \mathbf{M}_{N,j\ell}^2(\lambda_2)$. We obtain the set of AQS functions by adjusting the numerators of the concentrated quasi-scores with β and σ^2 being concentrated from (3.6):

$$S_{\mathtt{SPD}}^*(\lambda) = \begin{cases} \mathbb{Y}_N'(\lambda)\boldsymbol{\Phi}_1(\lambda)\mathbb{Y}_N(\lambda) - \frac{N}{N-p}\tilde{\sigma}_N^2(\lambda)\mathrm{tr}[\boldsymbol{\Phi}_1(\lambda)], \\ \mathbb{Y}_N'(\lambda)\boldsymbol{\Phi}_2(\lambda_2)\mathbb{Y}_N(\lambda) - \frac{N}{N-p}\tilde{\sigma}_N^2(\lambda)\mathrm{tr}[\boldsymbol{\Phi}_2(\lambda_2)], \\ \mathbf{Z}_N'\tilde{\boldsymbol{\zeta}}_N^*(\lambda), \end{cases} \tag{3.16}$$

where $\boldsymbol{\Phi}_1(\lambda) = \mathbf{M}_N(\lambda_2)[\mathbf{B}_{2N}(\lambda_2)\mathbf{G}_{1N}(\lambda_1)\mathbf{B}_{2N}^{-1}(\lambda_2) - \bar{\mathbf{G}}_{1N}(\lambda_1)I_N]$, $\boldsymbol{\Phi}_2(\lambda_2) = \mathbf{M}_N(\lambda_2)[\mathbf{G}_{2N}(\lambda_2) - \bar{\mathbf{G}}_{2N}(\lambda_2)I_N]\mathbf{M}_N(\lambda_2)$, $\bar{\mathbf{G}}_{rN}(\lambda_r) = \frac{1}{N}\mathrm{tr}[\mathbf{G}_{rN}(\lambda_r)], r = 1,2$, and $\mathbf{Z}_N = 1_{T-1}\otimes Z_n$. We see that $\mathrm{E}[S_{\mathtt{SPD}}^*(\lambda_0)|_{H_0}] = 0$, and that under H_0 at the true θ_0, the first two components of $S_{\mathtt{SPD}}^*(\lambda_0)$ can be written as $\mathbf{V}_N'\boldsymbol{\Phi}_1^*\mathbf{V}_N + \mathbf{V}_N'\Pi$ and $\mathbf{V}_N'\boldsymbol{\Phi}_2^*\mathbf{V}_N$, where $\Pi = \mathbf{M}_N[\mathbf{B}_{2N}\mathbf{G}_{1N}\mathbf{B}_{2N}^{-1} - \bar{\mathbf{G}}_{1N}I_N]\mathbb{X}_N(\lambda_{20})\beta_0$, and $\boldsymbol{\Phi}_r^* = \boldsymbol{\Phi}_r - \frac{1}{N-p}\mathrm{tr}(\boldsymbol{\Phi}_r)\mathbf{M}_N$, $r=1,2$. Define $\mathbf{g}_{Nj}^*(\theta_0), j=1,\ldots,N$, in the same way as $\mathbf{g}_{ni}^*(\theta_0)$ in (2.18), we have an AQS test for H_0 for the FE-SPD model:

$$T_{\mathtt{SPD}}^{\mathrm{r}*} = \big(\textstyle\sum_{j=1}^N \tilde{\mathbf{g}}_{Nj,\alpha}^{*\prime}\big)\big[\sum_{j=1}^N(\tilde{\mathbf{g}}_{Nj,\alpha}^* - \tilde{\Gamma}_N^*\tilde{\mathbf{g}}_{Nj,\lambda}^*)(\tilde{\mathbf{g}}_{Nj,\alpha}^* - \tilde{\Gamma}_N^*\tilde{\mathbf{g}}_{Nj,\lambda}^*)'\big]^{-1}\big(\sum_{j=1}^N \tilde{\mathbf{g}}_{Nj,\alpha}^*\big), \tag{3.17}$$

where $\tilde{\Gamma}_N^* = \tilde{\boldsymbol{\Sigma}}_{N,\alpha\lambda}^*\tilde{\boldsymbol{\Sigma}}_{N,\lambda\lambda}^{*-1}$, $\tilde{\boldsymbol{\Sigma}}_{N,\alpha\lambda}^* = -\frac{\partial}{\partial\lambda}S_{\mathtt{SPD},\alpha}^*(\tilde{\lambda}_N)$, and $\tilde{\boldsymbol{\Sigma}}_{N,\lambda\lambda}^* = -\frac{\partial}{\partial\lambda}S_{\mathtt{SPD},\lambda}^*(\tilde{\lambda}_N)$. These derivatives can be easily obtained from (3.16), and are given in Appendix C. Numerical derivatives may provide much simpler and yet quite accurate alternatives, as indicated in Footnote 9 for the SLR model. When the errors are normally distributed, one may simply use $\tilde{\boldsymbol{\Sigma}}_{N,\alpha\lambda}^* = \sum_{j=1}^N \tilde{\mathbf{g}}_{Nj,\alpha}^*\tilde{\mathbf{g}}_{Nj,\lambda}^{*\prime}$ and $\tilde{\boldsymbol{\Sigma}}_{N,\lambda\lambda}^* = \sum_{J=1}^N \tilde{\mathbf{g}}_{Nj,\lambda}^*\tilde{\mathbf{g}}_{Nj,\lambda}^{*\prime}$ based on an IME corresponding to an 'adjusted likelihood', leading to an adjusted score test, denoted by $T_{\mathtt{SPD}}^*$ for easy reference.[13]

[13]Alternatively, the generalized IME can be applied to give $\tilde{\boldsymbol{\Sigma}}_{N,\alpha\lambda}^* = \sum_{j=1}^N \tilde{\mathbf{g}}_{Nj,\alpha}^*\tilde{\mathbf{g}}_{Nj,\lambda}'$ and $\tilde{\boldsymbol{\Sigma}}_{N,\lambda\lambda}^* = \sum_{J=1}^N \tilde{\mathbf{g}}_{Nj,\lambda}^*\tilde{\mathbf{g}}_{Nj,\lambda}'$, where $\tilde{\mathbf{g}}_{Nj,\lambda}$ is the λ-component of $\tilde{\mathbf{g}}_{Nj}$ defined in (3.8), but Monte Carlo results show that the early version works better in finite samples.

Again, better finite sample properties of the tests based on the AS or QS functions can be achieved using the adjusted estimator: $\tilde{\lambda}_N^* = \arg\{S^*_{\mathtt{SPD},\lambda}(\lambda) = 0\}$ in place of the regular estimator $\tilde{\lambda}_N$ as it is typical that $S^*_{\mathtt{SPD},\lambda}(\tilde{\lambda}_N) \neq 0$.

Unlike the SLR model, the statistic $T^{\text{r}*}_{\mathtt{SPD}}$ may not be fully robust against non-normality. Similar to the developments leading to $T^{\mathtt{rr}}_{\mathtt{SPD}}$, a robust estimator of $\mathbf{\Omega}_N^* = \text{Var}[S^*_{\mathtt{SPD}}(\lambda_0)]$ is

$$\widetilde{\mathbf{\Omega}}^{\text{r}*}_N = \textstyle\sum_{j=1}^N(\tilde{\mathbf{g}}^*_{Nj}\tilde{\mathbf{g}}^{*\prime}_{Nj} + \tilde{\mathbf{d}}^*_{Nj}\tilde{\mathbf{d}}^{*\circ\prime}_{Nj}), \tag{3.18}$$

where $\mathbf{d}^*_{N,it} = \big\{(v^{*2}_{N,it} - \sigma_0^2)\phi^*_{1,it} + \Pi_{it}v^*_{it}, (v^{*2}_{N,it} - \sigma_0^2)\phi^*_{2,it}, z_{ni}\zeta_{N,it})\big\}$, and $\mathbf{d}^{*\circ}_{N,it} = \sum_{s(\neq t)=1}^{T-1}\mathbf{d}^*_{N,it}$. The coefficients $\phi^*_{r,it}$ represent the diagonal elements of $\mathbf{\Phi}^*_r$, $r = 1, 2$. Again $\text{Var}[S^*_{\mathtt{SPD},\alpha}(\tilde{\lambda}_N)]$ has two equivalent forms similar to (3.12) and (3.13). We take the latter and an AQS test statistic fully robust against non-normality takes a similar form as $T^{\mathtt{rr}}_{\mathtt{SPD}}$:

$$T^{\text{rr}*}_{\mathtt{SPD}} = \big(\textstyle\sum_{j=1}^N \tilde{\mathbf{g}}^{*\prime}_{Nj,\alpha}\big)\big[\sum_{j=1}^N(\tilde{\mathbf{s}}^*_{Nj}\tilde{\mathbf{s}}^{*\prime}_{Nj} + \tilde{\mathbf{f}}^*_{Nj}\tilde{\mathbf{f}}^{\circ*\prime}_{Nj})\big]^{-1}\big(\sum_{j=1}^N \tilde{\mathbf{g}}^*_{Nj,\alpha}\big), \tag{3.19}$$

where $\tilde{\mathbf{s}}^*_{Nj} = \tilde{\mathbf{g}}^*_{Nj,\alpha} - \tilde{\Gamma}^*_N\tilde{\mathbf{g}}^*_{Nj,\lambda}$, $\tilde{\mathbf{f}}^*_{Nj} = \tilde{\mathbf{d}}^*_{Nj,\alpha} - \tilde{\Gamma}^*_N\tilde{\mathbf{d}}^*_{Nj,\lambda}$, and $\tilde{\mathbf{f}}^{\circ*}_{Nj} = \tilde{\mathbf{d}}^{\circ*}_{Nj,\alpha} - \tilde{\Gamma}^*_N\tilde{\mathbf{d}}^{\circ*}_{Nj,\lambda}$; and $(\tilde{\mathbf{g}}^{*\prime}_{Nj,\lambda}, \tilde{\mathbf{g}}^{*\prime}_{Nj,\alpha})' = \tilde{\mathbf{g}}^*_{Nj}$, $(\tilde{\mathbf{d}}^{*\prime}_{Nj,\lambda}, \tilde{\mathbf{d}}^{*\prime}_{Nj,\alpha})' = \tilde{\mathbf{d}}^*_{Nj}$, and $(\tilde{\mathbf{d}}^{\circ*\prime}_{Nj,\lambda}, \tilde{\mathbf{d}}^{\circ*\prime}_{Nj,\alpha})' = \tilde{\mathbf{d}}^{\circ*}_{Nj}$.

Theorem 3.2. *Under the assumptions of Theorem 3.1, $T^{rr*}_{\mathtt{SPD}}|_{H_0} \xrightarrow{D} \chi^2_k$ when the errors are normal or non-normal; $T^*_{\mathtt{SPD}}|_{H_0} \xrightarrow{D} \chi^2_k$ when the errors are normal.*

3.1.4. Tests with temporal heterogeneity and heteroskedasticity

As a panel data model allows for a much richer structure than a cross-section model, we extend the above theory and method to a richer FE-SPD model. Besides allowing higher-order spatial lags in the response and the disturbance as indicated in Footnote 11, two immediate extensions are to allow time-specific FE and temporal heteroskedasticity, in addition to the individual-specific FE and the cross-sectional or spatial heteroskedasticity.

First, the **temporal heteroskedasticity** can be added to the FE-SPD model considered above by simply allowing $\{v_{it}\}$ to be independent $(0, \sigma^2 h(z'_{n,it}\alpha))$ with the values of the heteroskedasticity variables $z_{n,it}$ changing with both i and t. In this case, (3.4) becomes,

$$E(V^{*\prime}_{n1}, \ldots, V^{*\prime}_{n,T-1})'(V^{*\prime}_{n1}, \ldots, V^{*\prime}_{n,T-1}) = \sigma^2(F'_{T,T-1}\otimes I_n)\mathcal{H}_{nT}(\alpha)(F_{T,T-1}\otimes I_n) \equiv \sigma^2\mathbf{H}_N(\alpha),$$

where $\mathcal{H}_{nT}(\alpha) = \{h(z'_{n,it}\alpha)\}$. Thus, introducing time-wise heteroskedasticity induces time-wise non-zero correlation among $\{v^*_{it}\}$ although the cross-sectional independence is kept. Changes will occur in the expressions for the α-components of the score functions. However, there will be no additional technical complications as under the null, $\mathbf{H}_N(\alpha)|_{H_0} = I_N$

and $\{v_{it}\}$ become independent across both i and uncorrelated across t.

Further allowing **time-specific FE** in the FE-SPD model considered above gives it an added feature of being able to control (partially) for **temporal heterogeneity** as well. When the two types of FEs appear in the model additively, and when the spatial weight matrices are row-normalized, another layer of orthonormal transformation can be applied to wipe out the time FE (Lee and Yu, 2010; Yang et al., 2016). Let $F_{n,n-1}$ be the orthonormal eigenvector matrix of $J_n = I_n - \frac{1}{n} l_n l_n'$ corresponding to the eigenvalues of one. For $n \times 1$ vectors $A_{nt}, t = 1, \ldots, T$, where A_{nt} can be Y_{nt}, V_{nt}, and a column of X_{nt}, define

$$[A^*_{n-1,1}, \ldots, A^*_{n-1,T-1}] = F'_{n,n-1}[A_{n,1}, \ldots, A_{n,T}]F_{T,T-1},$$

and $W^*_{rn} = F'_{n,n-1} W_{rn} F_{n,n-1}$. Let $N = (n-1)(T-1)$ and define $\mathbf{Y}_N$, $\mathbf{X}_N$, $\mathbf{U}_N$ and $\mathbf{V}_N$ accordingly. Then, the transformed model takes an identical form as (3.3). We have, when heteroskedasticity exists along both cross-section and time dimensions,

$$E(\mathbf{V}_N \mathbf{V}_N') = \sigma^2 (F'_{T,T-1} \otimes F'_{n,n-1}) \mathcal{H}_{nT}(\alpha) (F_{T,T-1} \otimes F_{n,n-1}) \equiv \sigma^2 \mathbf{H}_N(\alpha).$$

Under the null we have $\mathbf{H}_N(\alpha)|_{H_0} = I_N$. Model estimation and the construction of the tests proceed as above. When the original errors are non-normal, additional complications will occur in the derivation of the QS tests, due to the lack of independence among the elements of $\mathbf{V}_N$ in both cross-section and time dimensions. For the same reason, proofs of the asymptotic properties of these tests will be more complicated as well. Along the line of (3.11), an extended OPMD estimator for the VC matrix of the QS function can be developed to give a QS test. The finite sample improved versions of the tests (AS and AQS) can be developed along the same line as well. Details are available upon request from the authors.

3.2. Homoskedasticity Tests: Extended FE-SPD Models

Major extensions to the FE-SPD model occur when one or more of the following features are allowed in the model: (*i*) time-varying spatial weight matrices, (*ii*) time-varying regression and spatial coefficients, (*iii*) unbalanced panels, and (*iv*) interactive fixed effects (IFE). As discussed in the introduction, adding any of these features in the FE-SPD model would render the conventional transformation method inapplicable in dealing with the *incidental parameters* – the FEs. Estimation of the structural parameters based on the concentrated (quasi) scores (with FEs being concentrated) would lead to inconsistent or asymptotically biased estimation and thereby inconsistent or asymptotically biased homoskedasticity tests. However, with proper adjustments to remove the effects of estimating the fixed effects, the adjusted (quasi) scores lead to asymptotically valid tests for homoskedasticity. We now present some critical discussion on these important extensions. Many developments are straightforward following the discussion, which are available upon request from the authors, but a full and rigorous study of these extensions can only be done through future research.

(i) Time-varying spatial weights. When the spatial weight matrices in Model (3.1) are allowed to change with t to give W_{1nt} and W_{2nt}, the transformation method cannot be applied to handle the individual FE, μ_n. In fact, after the transformation μ_n is wiped out but the 'spatial lag' structure is lost and a proper (quasi) likelihood function cannot be found.

In this case, we may start with the joint (quasi) Gaussian loglikelihood $\ell(\psi, \mu_n)$ of (ψ, μ_n), and concentrate out μ_n to give the concentrated (quasi) loglikelihood $\ell^c(\psi) = \ell(\psi, \tilde{\mu}_n(\psi))$, where $\tilde{\mu}_n(\psi)$ is the constrained estimate of μ_n, given ψ:

$$\tilde{\mu}_n(\psi) = \tfrac{1}{T}\textstyle\sum_{t=1}^{T}(B_{1nt}(\lambda_1)Y_{nt} - X_{nt}\beta),$$

where $B_{1nt}(\lambda_1) = I_n - \lambda_1 W_{1nt}$. Then, we obtain the concentrated (quasi) score $S^c(\psi) = \frac{\partial}{\partial\psi}\ell^c(\psi)$. It is easy to show that $\mathrm{E}[S^c(\psi_0)] = \big(0'_p, -\frac{n}{2\sigma_0^2}, -\frac{1}{T}\mathrm{tr}[\mathbf{B}_{1N}(\lambda_1)], -\frac{1}{T}\mathrm{tr}[\mathbf{B}_{2N}(\lambda_2)], 0'_k\big)'$, where $N = n \times T$, $\mathbf{B}_{1N}(\lambda_1) = \mathtt{blkdiag}\{B_{1nt}(\lambda_1)\}$, and similarly, $\mathbf{B}_{2N}(\lambda_2)$ is defined. Here, the operator $\mathtt{blkdiag}\{\cdots\}$ forms a block-diagonal matrix based on the given matrices. These give the AS or AQS function as $S^\circ(\psi_0) = S^c(\psi_0) - \mathrm{E}[S^c(\psi_0)]$, and the construction of the AS and AQS tests, for testing H_0: $\alpha = 0$, proceeds with $S^\circ(\psi)|_{H_0}$. The OPMD estimate of $\mathrm{Var}[S^\circ(\psi_0)|_{H_0}]$ is obtained in a similar way as (3.11). If necessary, the improved versions of the AS and AQS tests can be obtained by further concentrating $S^\circ(\psi)$ with respect to β and σ^2, and then re-adjusting in a similar manner.

(ii) Time-varying parameters. In the SPD model, we allow the regression coefficients and the spatial lag coefficient to vary with time, i.e., in Model (3.1) β is replaced by β_t and λ_1 by λ_{1t}. The spatial weight matrices can be time-varying as well. Clearly, with these extensions the conventional transformation method cannot be applied. Denote $\boldsymbol{\beta} = (\beta'_1, \ldots, \beta'_T)'$, $\boldsymbol{\lambda}_1 = (\lambda_{11}, \ldots, \lambda_{1T})'$, $\boldsymbol{\lambda} = (\boldsymbol{\lambda}'_1, \lambda_2)'$, and $\boldsymbol{\psi} = (\boldsymbol{\beta}', \sigma^2, \boldsymbol{\lambda}', \alpha')'$. Redefine $B_{1nt}(\lambda_{1t}) = I_n - \lambda_{1t}W_{1nt}, t = 1, \ldots, T$, and $\mathbf{B}_{1N}(\lambda_{1t}) = \mathtt{blkdiag}\{B_{1nt}(\lambda_{1t})\}$, with N, $B_{2nt}(\lambda_2)$ and $\mathbf{B}_{2N}(\lambda_2)$ being unchanged. We obtain the joint (quasi) Gaussian loglikelihood $\ell(\boldsymbol{\psi}, \mu_n)$ of $(\boldsymbol{\psi}, \mu_n)$, and the concentrated (quasi) loglikelihood $\ell^c(\boldsymbol{\psi}) = \ell(\boldsymbol{\psi}, \tilde{\mu}_n(\boldsymbol{\psi}))$, where

$$\tilde{\mu}_n(\boldsymbol{\psi}) = \tfrac{1}{T}\textstyle\sum_{t=1}^{T}(B_{1nt}(\lambda_{1t})Y_{nt} - X_{nt}\beta_t),$$

and the concentrated (quasi) score $S^c(\boldsymbol{\psi}) = \frac{\partial}{\partial \boldsymbol{\psi}} \ell^c(\boldsymbol{\psi})$. One can show that

$$\mathrm{E}[S^c(\boldsymbol{\psi}_0)] = \left[0'_{pT}, -\frac{n}{2\sigma_0^2}, -\frac{1}{T}\{\mathrm{tr}(B_{1nt}(\lambda_{1t})), t = 1, \ldots, T\}, -\frac{1}{T}\mathrm{tr}(\mathbf{B}_{2N}(\lambda_2)), 0'_k\right]'.$$

These give the AS or AQS function as $S^\circ(\boldsymbol{\psi}_0) = S^c(\boldsymbol{\psi}_0) - \mathrm{E}[S^c(\boldsymbol{\psi}_0)]$, and the construction of the AS and AQS tests, for testing H_0: $\alpha = 0$, proceeds with $S^\circ(\boldsymbol{\psi})|_{H_0}$. The OPMD estimate of $\mathrm{Var}[S^\circ(\boldsymbol{\psi}_0)|_{H_0}]$ is obtained in a similar way as (3.11). If necessary, the improved versions of the AS and AQS tests can be derived by further concentrating $S^\circ(\boldsymbol{\psi})$ with respect to $\boldsymbol{\beta}$ and σ^2, and then re-adjusting in a similar way.

(iii) Unbalanced panels. Suppose only n_t out of n spatial units appeared in period t. Let D_t be a selection matrix obtained from the $n \times n$ identity matrix I_n by deleting its rows corresponding to the 'missing' spatial units in period t. The FE-SPD model becomes:

$$Y_{nt} = \lambda_1 W_{1nt} Y_{nt} + X_{nt}\beta + D_t \mu_n + U_{nt}, \;\; U_{nt} = \lambda_2 W_{2nt} U_{nt} + V_{nt}, \;\; t = 1, 2, \ldots, T. \quad (3.20)$$

where Y_{nt} is $n_t \times 1$, $W_{rnt}, r = 1, 2$, are $n_t \times n_t$, *etc.* Thus, W_{rnt} cannot be time-invariant.

With an unbalanced panel, we are facing a similar problem as the case of time-varying spatial weights since no transformation can be used to eliminate the FE. Under a random 'missing' scheme (randomly dropping out and randomly joining in), our AQS method may again be able to provide a feasible solution to the testing problem for this model by adjusting the concentrated (quasi) scores to account for the direct estimation of the fixed effects.

Let $N = \sum_1^T n_t$. Define $\mathbf{W}_{rN} = \mathtt{blkdiag}\{W_{rn1}, \ldots, W_{rnT}\}, r = 1, 2$. The model can be written in matrix form: $\mathbf{Y}_N = \lambda_1 \mathbf{W}_{1N} \mathbf{Y}_N + \mathbf{X}_N \beta + \mathbf{D}_N \mu_n + \mathbf{U}_N$, $\mathbf{U}_N = \lambda_2 \mathbf{W}_{2N} \mathbf{U}_N + \mathbf{V}_N$. The joint (quasi) Gaussian loglikelihood $\ell(\psi, \mu_n)$ and its concentrated version $\ell^c(\psi) = \ell(\psi, \tilde{\mu}_n(\psi))$ can be easily obtained, where

$$\tilde{\mu}_n(\psi) = [\mathbb{B}'_{2N}(\lambda_2)\mathbb{B}_{2N}(\lambda_2)]^{-1}\mathbb{B}'_{2N}(\lambda_2)\mathbf{B}_{2N}(\lambda_2)[\mathbf{B}_{1N}(\lambda_1)\mathbf{Y}_N - \mathbf{X}_N\beta],$$

with $\mathbb{B}_{2N}(\lambda_2) = \mathbf{B}_{2N}(\lambda_2)\mathbf{D}_N$, and $\mathbf{B}_{rN}(\lambda_r) = I_N - \lambda_r \mathbf{W}_{rN}, r = 1, 2$. Along similar lines, one can obtain the AQS function to remove the effect of estimating μ_n, derive the (extended) OPMD estimate of its VC matrix, and obtain the desired tests.

(iv) Interactive fixed effects. The unit-specific and the time-specific fixed effects may appear in Model (3.1) interactively, i.e., $\mu_n a_t$, where a_t can be an $r \times 1$ vector and μ_n an $n \times r$ matrix. This gives rise to an SPD model with interactive fixed effects or multi-factor structure. In this case, the transformation method again cannot be used to deal with these incidental parameters, even though the other aspects of the standard FE-SPD model remain, unless $r = 1$. With the spatial weights varying over time, the transformation approach does not offer a solution to our testing problem even when $r = 1$. The proposed AQS method offers a feasible solution, at least when T is fixed. Let $\ell(\psi, \mu_n, \mathbf{a}_T)$ be the joint (quasi) Gaussian loglikelihood of $(\psi, \mu_n, \mathbf{a}_T)$, where $\mathbf{a}_T = (a_1, \ldots, a_T)'$. Consider time-varying spatial weight matrices $W_{rnt}, r = 1, 2$. When $r = 1$, the results in case **(i)** can be extended upon dividing a_t on both sides of the model, i.e., given $(\psi, \mathbf{a}_T)$, $\ell(\psi, \mu_n, \mathbf{a}_T)$ is partially maximized at

$$\tilde{\mu}_n(\psi, \mathbf{a}_T) = \tfrac{1}{T}\sum_{t=1}^{T} \tfrac{1}{a_t}(B_{1nt}(\lambda_1)Y_{nt} - X_{nt}\beta),$$

which gives the concentrated (quasi) loglikelihood $\ell^c(\psi, \mathbf{a}_T) = \ell(\psi, \tilde{\mu}_n(\psi, \mathbf{a}_T), \mathbf{a}_T)$. The rest follows the AQS idea with proper constraints imposed on $\mathbf{a}_T$ to ensure parameter identifiability. When $r > 1$, this simple method does not apply, and a general solution is:

$$\tilde{\mu}_n(\psi, \mathbf{a}_T) = \big\{B_{1nt}(\lambda_1)Y_{nt} - X_{nt}\beta, t = 1, \ldots, T\big\}\mathbf{a}_T(\mathbf{a}_T'\mathbf{a}_T)^{-1}.$$

After a proper reparameterization on the $T \times r$ matrix $\mathbf{a}_T$ to ensure parameter identifiability (see, Bai and Ng, 2013), the same ideas may lead to the desired AQS function, and thus the desired tests for homoskedasticity.

4. Tests for Homoskedasticity: Dynamic SPD Model

As discussed in the introduction, the AQS-OPMD idea goes much beyond merely providing a simpler method of constructing tests for homoskedasticity and its finite sample improved version in standard scenarios where the conventional methods are available – it provides feasible solutions to "non-standard problems" where the usual methods fail due to the lack of (i) a valid (quasi) score and (ii) a feasible method for VC matrix estimation. The **dynamic** spatial panel data (DSPD) model with short panels provides a perfect example of this. In this case, even if all the requirements as for a standard SPD model are met, one is still unable to achieve either (i) or (ii) due to the well-known *initial values problem* (IVP), as argued in the introduction. The conditional (quasi) score, treating the initial values as exogenously given, may be the best we can get, but it does not lead to consistent estimation when T is fixed and it incurs an asymptotic bias when T goes large with n (Yu et al., 2008). Again, this conditional quasi score can be adjusted to 'remove' the effect of IVP and to give an AQS vector for the heteroskedastic DSPD model that is unbiased and consistent. Furthermore, the OPMD method provides a feasible estimate of the VC matrix of the AQS vector – together the AQS and OPMD methods lead to asymptotically valid tests for homoskedasticity.

4.1. The Dynamic FE-SPD Model and its M-Estimation

The dynamic SPD (DSPD) model with fixed effects, hereafter FE-DSPD, takes the following form after first-differencing to eliminate the individual-specific fixed effects:

$$\Delta Y_{nt} = \rho\Delta Y_{n,t-1} + \lambda_1 W_{1n}\Delta Y_{nt} + \lambda_2 W_{2n}\Delta Y_{n,t-1} + \Delta X_{nt}\beta + \Delta U_{nt}, \tag{4.1}$$
$$\Delta U_{nt} = \lambda_3 W_{3n}\Delta U_{nt} + \Delta V_{nt}, \quad t = 2, \ldots, T.$$

It extends the FE-SPD model by adding the dynamic term $\rho Y_{n,t-1}$ and the space-time lag term $\lambda_2 W_{2n} Y_{n,t-1}$, and the FE-DSPD model considered in Yang (2018a) by allowing for cross-sectional heteroskedasticity, i.e., $V_{nt} \sim (0, \sigma^2\mathcal{H}_n(\alpha)), t = 1, \ldots, T$. It also allows the time-specific effects and the spatial Durbin effects as in Yang (2018a), of which both appear in the model in the form of additional regressors embedded in ΔX_{nt}. A test for cross-sectional homoskedasticity again corresponds to the test of the null hypothesis $H_0 : \alpha = 0$.

Stacking the vectors and matrices in (4.1) for $t = 2, \ldots, T$, i.e., $\Delta\mathbf{Y}_N = \{\Delta Y'_{n2}, \ldots, \Delta Y'_{nT}\}'$, $\Delta\mathbf{Y}_{N,-1} = \{\Delta Y'_{n1}, \ldots, \Delta Y'_{n,T-1}\}'$, and similarly for $\Delta\mathbf{X}_N$ and $\Delta\mathbf{V}_N$, where $N = n(T-1)$. Let $\mathbf{W}_{rN} = I_{T-1} \otimes W_{rn}, r = 1, 2, 3$, $\mathbf{B}_{rN}(\lambda_r) = I_{T-1} \otimes B_{rn}(\lambda_r), r = 1, 3$, and $\mathbf{B}_{2N}(\rho, \lambda_2) = I_{T-1} \otimes B_{2n}(\rho, \lambda_2)$, where $B_{rn}(\lambda_r) = I_n - \lambda_r W_{rn}, r = 1, 3$, and $B_{2n}(\rho, \lambda_2) = \rho I_n + \lambda_2 W_{2n}$. Let $\mathcal{Z}_{nj}$ be the diagonal matrix formed by the jth column Z_{nj} of Z_n, where Z_n is the $n \times k$ matrix of the k heteroskedasticity variables. Denote $\lambda = (\lambda_1, \lambda_2, \lambda_3)'$, $\delta = (\rho, \lambda')'$, and $\theta = (\beta', \sigma^2, \delta')'$. Define $\mathbf{C}_N = C_{T-1} \otimes I_n$, where C_{T-1} is a $(T-1) \times (T-1)$ constant matrix with 2 on the main diagonal and -1 on the two parallel diagonals and 0 otherwise.

If V_{nt} are independent $N(0, \sigma^2\mathcal{H}_n(\alpha))$, then $\Delta\mathbf{V}_N \sim N[0, \sigma^2 C_{T-1} \otimes \mathcal{H}_n(\alpha)]$. From this, one can easily obtain the conditional (quasi) Gaussian loglikelihood $\ell(\theta, \alpha)$ for (θ, α), given ΔY_{n1}, and the conditional (quasi) Gaussian score function $S_{\texttt{DSPD}}(\theta, \alpha) = \frac{\partial}{\partial(\theta', \alpha')'}\ell(\theta, \alpha)$. Assume,

Initial conditions (Assumption A, Yang, 2018a): *(i) the processes started m periods before the start of data collection, the 0th period, and (ii) if $m \geq 1$, ΔY_{n0} is independent of future errors $\{V_{nt}, t \geq 1\}$; if $m = 0$, Y_{n0} is independent of future errors $\{V_{nt}, t \geq 1\}$.*

Under the above very minimum knowledge about the processes in the past, and assuming further that $B_{1n}^{-1}(\lambda_{10})$ and $B_{3n}^{-1}(\lambda_{30})$ exist, one derives the analytical expression for $\mathrm{E}[S_{\texttt{DSPD}}(\theta_0, 0)]$ by following the method in Yang (2018a), which is a function of only the common parameters θ and the observables and has the $(\rho, \lambda_1, \lambda_2)$-elements being non-zero and of order $O(n)$. This shows that $S_{\texttt{DSPD}}(\theta_0, 0)$ needs to be corrected, and that it can be corrected as $S^{\circ}_{\texttt{DSPD}}(\theta_0) = S_{\texttt{DSPD}}(\theta_0, 0) - \mathrm{E}[S_{\texttt{DSPD}}(\theta_0, 0)]$, leading to the desired AQS function at the null:

$$S^{\circ}_{\mathtt{DSPD}}(\theta)=\begin{cases}\frac{1}{\sigma^2}\Delta\mathbf{X}'_N\mathbf{C}'^{-1}_N\mathbf{B}'_{3N}(\lambda_3)\Delta\mathbf{V}_N(\beta,\delta),\\ \frac{1}{2\sigma^4}\Delta\mathbf{V}'_N(\beta,\delta)\mathbf{C}^{-1}_N\Delta\mathbf{V}_N(\beta,\delta)-\frac{N}{2\sigma^2},\\ \frac{1}{\sigma^2}\Delta\mathbf{V}'_N(\beta,\delta)\mathbf{C}^{-1}_N\mathbf{B}_{3N}(\lambda_3)\Delta\mathbf{Y}_{N,-1}+\mathrm{tr}(\mathbf{C}^{-1}_N\mathbf{D}_{N,-1}),\\ \frac{1}{\sigma^2}\Delta\mathbf{V}'_N(\beta,\delta)\mathbf{C}^{-1}_N\mathbf{B}_{3N}(\lambda_3)\mathbf{W}_{1N}\Delta\mathbf{Y}_N+\mathrm{tr}(\mathbf{C}^{-1}_N\mathbf{D}_N\mathbf{W}_{1N}),\\ \frac{1}{\sigma^2}\Delta\mathbf{V}'_N(\beta,\delta)\mathbf{C}^{-1}_N\mathbf{B}_{3N}(\lambda_3)\mathbf{W}_{2N}\Delta\mathbf{Y}_{N,-1}+\mathrm{tr}(\mathbf{C}^{-1}_N\mathbf{D}_{N,-1}\mathbf{W}_{2N}),\\ \frac{1}{\sigma^2}\Delta\mathbf{V}'_N(\beta,\delta)(C^{-1}_{T-1}\otimes G_{3n}(\lambda_3))\Delta\mathbf{V}_N(\beta,\delta)-(T-1)\mathrm{tr}(G_{3n}(\lambda_3)),\\ \frac{1}{2\sigma^2}\Delta\mathbf{V}'_N(\beta,\delta)(C^{-1}_{T-1}\otimes\mathcal{Z}_{nj})\Delta\mathbf{V}_N(\beta,\delta)-\frac{(T-1)}{2}Z'_{nj}1_n,\ j=1,\ldots,k,\end{cases}\tag{4.2}$$

where $\Delta\mathbf{V}_N(\beta,\delta)=\mathbf{B}_{3N}(\lambda_3)[\mathbf{B}_{1N}(\lambda_1)\Delta\mathbf{Y}_N-\mathbf{B}_{2N}(\rho,\lambda_2)\Delta\mathbf{Y}_{N,-1}-\Delta\mathbf{X}_N\beta]$,

$$\mathbf{D}_{N,-1}=\begin{pmatrix}I_n, & 0, & \ldots & 0, & 0\\ \mathcal{B}_n-2I_n, & I_n, & \ldots & 0, & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \mathcal{B}_n^{T-4}(I_n-\mathcal{B}_n)^2, & \mathcal{B}_n^{T-5}(I_n-\mathcal{B}_n)^2, & \ldots & \mathcal{B}_n-2I_n, & I_n\end{pmatrix}\mathbf{B}^{-1}_{1N}(\lambda_1),$$

$$\text{and }\mathbf{D}_N=\begin{pmatrix}\mathcal{B}_n-2I_n, & I_n, & \ldots & 0\\ (I_n-\mathcal{B}_n)^2, & \mathcal{B}_n-2I_n, & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ \mathcal{B}_n^{T-3}(I_n-\mathcal{B}_n)^2, & \mathcal{B}_n^{T-4}(I_n-\mathcal{B}_n)^2, & \ldots & \mathcal{B}_n-2I_n\end{pmatrix}\mathbf{B}^{-1}_{1N}(\lambda_1),$$

$G_{3n}(\lambda_3)=W_{3n}B^{-1}_{3n}(\lambda_3)$ and $\mathcal{B}_n\equiv\mathcal{B}_n(\rho,\lambda_1,\lambda_2)=B^{-1}_{1n}(\lambda_1)B_{2n}(\rho,\lambda_2)$. Let $S^{\circ}_{\mathtt{DSPD},\theta}(\theta)$ and $S^{\circ}_{\mathtt{DSPD},\alpha}(\theta)$ be, respectively, the θ- and α-component of $S^{\circ}_{\mathtt{DSPD}}(\theta)$. The solution of $S^{\circ}_{\mathtt{DSPD},\theta}(\theta)=0$ gives the M-estimator $\tilde\theta_{\mathtt{M}}$ of θ in the null model, and its asymptotic properties are given in Yang (2018a). The unknown constant $\dot h(0)$ in $S^{\circ}_{\mathtt{DSPD},\alpha}(\theta)$ is dropped for the same reason.

4.2. The AQS Test for Homoskedasticity

Note that $\mathcal{B}_n$ and hence $\mathbf{D}_{N,-1}$ and $\mathbf{D}_N$ depend on $(\rho,\lambda_1,\lambda_2)$, and are denoted in the same way at the true parameter values $(\rho_0,\lambda_{10},\lambda_{20})$. For other parametric quantities at the true parameter values, shorthand notation will again be used, e.g., $\mathbf{B}_{1N}$ for $\mathbf{B}_{1N}(\lambda_{10})$ and G_{3n} for $G_{3n}(\lambda_{30})$. Now, from Lemma 3.2 of Yang (2018a), we have

$$\Delta\mathbf{Y}_N=\mathbb{R}\ \Delta\mathbf{Y}_{N1}+\boldsymbol{\eta}+\mathbb{S}\Delta\mathbf{V}_N,\quad\text{and}\quad\Delta\mathbf{Y}_{N,-1}=\mathbb{R}_{-1}\Delta\mathbf{Y}_{N1}+\boldsymbol{\eta}_{-1}+\mathbb{S}_{-1}\Delta\mathbf{V}_N,\tag{4.3}$$

where $\Delta\mathbf{Y}_{N1}=1_{T-1}\otimes\Delta Y_{n1}$, $\mathbb{R}=\mathrm{blkdiag}(\mathcal{B}_n,\mathcal{B}_n^2,\ldots,\mathcal{B}_n^{T-1})$, $\mathbb{R}_{-1}=\mathrm{blkdiag}(I_n,\mathcal{B}_n,\ldots,\mathcal{B}_n^{T-2})$, $\boldsymbol{\eta}=\mathbb{B}\mathbf{B}^{-1}_{1N}\Delta\mathbf{X}_N\beta_0$, $\boldsymbol{\eta}_{-1}=\mathbb{B}_{-1}\mathbf{B}^{-1}_{1N}\Delta\mathbf{X}_N\beta_0$, $\mathbb{S}=\mathbb{B}\mathbf{B}^{-1}_{1N}\mathbf{B}^{-1}_{3N}$, $\mathbb{S}_{-1}=\mathbb{B}_{-1}\mathbf{B}^{-1}_{1N}\mathbf{B}^{-1}_{3N}$,

$$\mathbb{B}=\begin{pmatrix}I_n & 0 & \ldots & 0 & 0\\ \mathcal{B}_n & I_n & \ldots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \mathcal{B}_n^{T-2} & \mathcal{B}_n^{T-3} & \ldots & \mathcal{B}_n & I_n\end{pmatrix},\quad\textit{and}\quad\mathbb{B}_{-1}=\begin{pmatrix}0 & 0 & \ldots & 0 & 0\\ I_n & 0 & \ldots & 0 & 0\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \mathcal{B}_n^{T-3} & \mathcal{B}_n^{T-4} & \ldots & I_n & 0\end{pmatrix}.$$

From (4.3), one immediately obtains,

$$S^{\circ}_{\texttt{DSPD}}(\theta_0) = \begin{cases} \Pi_1' \Delta \mathbf{V}_N, \\ \Delta \mathbf{V}_N' \mathbf{\Phi}_1 \Delta \mathbf{V}_N - \frac{N}{2\sigma^2}, \\ \Delta \mathbf{V}_N' \mathbf{\Psi}_1 \Delta \mathbf{Y}_{N1} + \Pi_2' \Delta \mathbf{V}_N + \Delta \mathbf{V}_N' \mathbf{\Phi}_2 \Delta \mathbf{V}_N + \mathrm{tr}(\mathbf{C}_N^{-1} \mathbf{D}_{N,-1}), \\ \Delta \mathbf{V}_N' \mathbf{\Psi}_2 \Delta \mathbf{Y}_{N1} + \Pi_3' \Delta \mathbf{V}_N + \Delta \mathbf{V}_N' \mathbf{\Phi}_3 \Delta \mathbf{V}_N + \mathrm{tr}(\mathbf{C}_N^{-1} \mathbf{D}_N \mathbf{W}_{1N}), \\ \Delta \mathbf{V}_N' \mathbf{\Psi}_3 \Delta \mathbf{Y}_{N1} + \Pi_4' \Delta \mathbf{V}_N + \Delta \mathbf{V}_N' \mathbf{\Phi}_4 \Delta \mathbf{V}_N + \mathrm{tr}(\mathbf{C}_N^{-1} \mathbf{D}_{N,-1} \mathbf{W}_{2N}), \\ \Delta \mathbf{V}_N' \mathbf{\Phi}_5 \Delta \mathbf{V}_N - (T-1)\mathrm{tr}(G_{3n}), \\ \Delta \mathbf{V}_N' \mathbf{\Phi}_{5+j} \Delta \mathbf{V}_N - \frac{(T-1)}{2} Z_{nj}' 1_n, \quad j = 1, \dots, k, \end{cases} \tag{4.4}$$

where $\mathbf{\Phi}_{5+j} = \frac{1}{2\sigma_0^2}(C_{T-1}^{-1} \otimes \mathcal{Z}_{nj}), \quad j = 1, \dots, k,$

$$\begin{array}{lll} \Pi_1 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\Delta\mathbf{X}_N, & \mathbf{\Psi}_1 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\mathbb{R}_{-1}, & \mathbf{\Phi}_2 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\mathbb{S}_{-1} \\ \Pi_2 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\boldsymbol{\eta}_{-1}, & \mathbf{\Psi}_2 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\mathbf{W}_{1N}\mathbb{R}, & \mathbf{\Phi}_3 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\mathbf{W}_{1N}\mathbb{S} \\ \Pi_3 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\mathbf{W}_{1N}\boldsymbol{\eta}, & \mathbf{\Psi}_3 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\mathbf{W}_{2N}\mathbb{R}_{-1}, & \mathbf{\Phi}_4 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\mathbf{W}_{2N}\mathbb{S}_{-1} \\ \Pi_4 = \frac{1}{\sigma_0^2}\mathbf{C}_N^{-1}\mathbf{B}_{3N}\mathbf{W}_{2N}\boldsymbol{\eta}_{-1}, & \mathbf{\Phi}_1 = \frac{1}{2\sigma_0^4}\mathbf{C}_N^{-1}, & \mathbf{\Phi}_5 = \frac{1}{\sigma_0^2}(C_{T-1}^{-1} \otimes G_{3n}) \end{array}$$

The expression for $S^{\circ}_{\texttt{DSPD}}(\theta_0)$ given in (4.4) shows clearly that the usual plug-in method for estimating $\mathbf{\Sigma}_N = \mathrm{Var}[S^{\circ}_{\texttt{DSPD}}(\theta_0)]$ does not work as the analytical expression of $\mathbf{\Sigma}_N$ involves the unobservables contained in ΔY_{n1}, $\boldsymbol{\eta}_{-1}$ and $\boldsymbol{\eta}$. We show that an OPMD estimate of $\mathbf{\Sigma}_N$ can be derived when T is fixed, following the methods of Yang (2018a).

Now, for the general matrices Π, $\mathbf{\Phi}$ and $\mathbf{\Psi}$ appearing in (4.4), denote by Π_t, Φ_{ts} and Ψ_{ts} their submatrices partitioned according to $t, s = 2, \dots, T$. Define $\Psi_{t+} = \sum_{s=2}^{T} \Psi_{ts}$, $t = 2, \dots, T$, $\Theta = \Psi_{2+}(B_{30}B_{10})^{-1}$, $\Delta Y_{n1}^{\circ} = B_{30}B_{10}\Delta Y_{n1}$, and $\Delta Y_{n1t}^{*} = \Psi_{t+}\Delta Y_{n1}$. Let $\{\mathcal{G}_{n,i}\}$ be the increasing sequence of σ-fields generated by $(v_{j1}, \dots, v_{jT}, j = 1, \dots, i), i = 1, \dots, n, n \geq 1$. Let $\mathcal{F}_{n,0}$ be the σ-field generated by $(v_0, \Delta y_0)$, and define $\mathcal{F}_{n,i} = \mathcal{F}_{n,0} \otimes \mathcal{G}_{n,i}$. Clearly, $\mathcal{F}_{n,i-1} \subseteq \mathcal{F}_{n,i}$, i.e., $\{\mathcal{F}_{n,i}\}_{i=1}^{n}$ is an increasing sequence of σ-fields, for each $n \geq 1$. Using Lemma 3.3 of Yang (2018a), the typical terms in (4.4) can be written as $\Pi' \Delta \mathbf{V}_N = \sum_{i=1}^{n} g_{1i}$, $\Delta \mathbf{V}_N' \mathbf{\Phi} \Delta \mathbf{V}_N - \mathrm{E}(\Delta \mathbf{V}_N' \mathbf{\Phi} \Delta \mathbf{V}_N) = \sum_{i=1}^{n} g_{2i}$, and $\Delta \mathbf{V}_N' \mathbf{\Psi} \Delta \mathbf{Y}_{N1} - \mathrm{E}(\Delta \mathbf{V}_N' \mathbf{\Psi} \Delta \mathbf{Y}_{N1}) = \sum_{i=1}^{n} g_{3i}$, so that $\{(g_{1i}', \ g_{2i}, \ g_{3i})', \mathcal{F}_{n,i}\}_{i=1}^{n}$ form a vector MD sequence, where

$$g_{1i} = \sum_{t=2}^{T} \Pi_{it}' \Delta v_{it}, \tag{4.5}$$

$$g_{2i} = \sum_{t=2}^{T} (\Delta v_{it} \Delta \xi_{it} + \Delta v_{it} \Delta v_{it}^{*} - \sigma_{v0}^2 d_{it}), \tag{4.6}$$

$$g_{3i} = \Delta v_{2i} \Delta \zeta_i + \Theta_{ii}(\Delta v_{2i} \Delta y_{1i}^{\circ} + \sigma_{v0}^2) + \sum_{t=3}^{T} \Delta v_{it} \Delta y_{1it}^{*}, \tag{4.7}$$

$\{\Delta \xi_{it}\} = \Delta \xi_t = \sum_{s=2}^{T} (\Phi_{st}^{u\prime} + \Phi_{ts}^{l}) \Delta V_{ns}$, $\Delta V_{nt}^{*} = \sum_{s=2}^{T} \Phi_{ts}^{d} \Delta V_{ns}$, $\{d_{it}\}$ = diagonal elements of $\mathbf{C}_N \mathbf{\Phi}$, $\{\Delta \zeta_i\} = \Delta \zeta = (\Theta^u + \Theta^l) \Delta Y_{n1}^{\circ}$, $\Delta Y_{n1}^{\circ} = B_{3n} B_{1n} \Delta Y_{n1}$, and $\mathrm{diag}\{\Theta_{ii}\} = \Theta^d$.

Finally, for each $\Pi_r, r = 1, 2, 3, 4$, appearing in (4.4), define g_{1ri} according to (4.5). For each $\Phi_r, r = 1, \ldots, 5+k$, appearing in (4.4), define g_{2ri} according to (4.6). For each $\Psi_r, r = 1, 2, 3$, appearing in (4.4), define g_{3ri} according to (4.7). Let

$$\mathbf{g}_{ni} = (g'_{11i},\ g_{21i},\ g_{31i} + g_{12i} + g_{22i},\ g_{32i} + g_{13i} + g_{23i},\ g_{33i} + g_{14i} + g_{24i},\ g_{25i}, \ldots, g_{2(5+k)i})'.$$

Then, $S^{\circ}_{\mathtt{DSPD}}(\theta_0) = \sum_{i=1}^{n} \mathbf{g}_{ni}$, and $\{\mathbf{g}_{ni}, \mathcal{F}_{n,i}\}$ form a vector MD sequence. It follows that $\mathbf{\Omega}_N = \text{Var}[S^{\circ}_{\mathtt{DSPD}}(\theta_0)] = \sum_{i=1}^{n} \text{E}(\mathbf{g}_{ni}\mathbf{g}'_{ni})$. The 'average' of the outer products of the estimated $\mathbf{g}'_{ni}s$ at H_0, i.e., $\frac{1}{N}\sum_{i=1}^{n} \tilde{\mathbf{g}}_{ni}\tilde{\mathbf{g}}'_{ni}$, gives a consistent estimate of $\frac{1}{N}\mathbf{\Omega}_N$. Partition $S^{\circ}_{\mathtt{DSPD}}(\theta)$, $\mathbf{g}_{ni}$ and $\mathbf{\Omega}_N$ according θ and α, and denote the subvectors and submatrices by adding subscripts θ and/or α, as in Section 3. Let $\mathbf{\Sigma}_{N,\theta\theta} = -\text{E}[\frac{\partial}{\partial\theta} S^{\circ}_{\mathtt{DSPD},\theta}(\theta_0)]$ and $\mathbf{\Sigma}_{N,\alpha\theta} = -\text{E}[\frac{\partial}{\partial\theta} S^{\circ}_{\mathtt{DSPD},\alpha}(\theta_0)]$. For the same reasoning, an asymptotic MD representation is developed for $S^{\circ}_{\mathtt{DSPD},\alpha}(\tilde{\theta}_{\mathtt{M}})$, and an AQS test statistic robust against nonnormality is obtained:

$$T^{\mathrm{r}}_{\mathtt{DSPD}} = \big(\textstyle\sum_{i=1}^{n} \tilde{\mathbf{g}}'_{ni,\alpha}\big)\big[\sum_{i=1}^{n}(\tilde{\mathbf{g}}_{ni,\alpha} - \widetilde{\mathbf{\Gamma}}_N\tilde{\mathbf{g}}_{ni,\theta})(\tilde{\mathbf{g}}_{ni,\alpha} - \widetilde{\mathbf{\Gamma}}_N\tilde{\mathbf{g}}_{ni,\theta})'\big]^{-1}\big(\sum_{j=1}^{n} \tilde{\mathbf{g}}_{ni,\alpha}\big), \tag{4.8}$$

where $\widetilde{\mathbf{\Gamma}}_N = \widetilde{\mathbf{\Sigma}}_{N,\alpha\theta}\widetilde{\mathbf{\Sigma}}^{-1}_{N,\theta\theta}$, $\widetilde{\mathbf{\Sigma}}_{N,\alpha\theta} = -\frac{\partial}{\partial\theta'} S^{\circ}_{\mathtt{DSPD},\alpha}(\tilde{\theta}_{\mathtt{M}})$, $\widetilde{\mathbf{\Sigma}}_{N,\theta\theta} = -\frac{\partial}{\partial\theta'} S^{\circ}_{\mathtt{DSPD},\theta}(\tilde{\theta}_{\mathtt{M}})$, and $\tilde{\theta}_{\mathtt{M}}$ is the M-estimator of Yang (2018a) for the null model, which solves $S^{\circ}_{\mathtt{DSPD},\theta}(\theta) = 0$. These derivatives can be easily obtained from (4.2) and are given in Appendix D.[14]

Theorem 4.1. *Extending Assumption 2.2 to $\Delta\mathbf{X}_N$ and assuming $\frac{1}{N}\mathbf{\Omega}_N$ is p.d. for large enough N, under the regularity conditions of Yang (2018a) we have under H_0, $T^{r}_{\mathtt{DSPD}} \xrightarrow{D} \chi^2_k$.*

4.3. Finite Sample Improved AQS Test

To improve the finite sample performance, the test given above can be further adjusted by working with the AQS function concentrating β and σ^2 from (4.2):

$$S^{\circ c}_{\mathtt{DSPD}}(\delta) = \begin{cases} \frac{1}{\tilde{\sigma}^2_N(\delta)}\Delta\widetilde{\mathbf{V}}'_N(\delta)\mathbf{C}_N^{-1}\mathbf{B}_{3N}(\lambda_3)\Delta\mathbf{Y}_{N,-1} + \text{tr}(\mathbf{C}_N^{-1}\mathbf{D}_{N,-1}), \\ \frac{1}{\tilde{\sigma}^2_N(\delta)}\Delta\widetilde{\mathbf{V}}'_N(\delta)\mathbf{C}_N^{-1}\mathbf{B}_{3N}(\lambda_3)\mathbf{W}_{1N}\Delta\mathbf{Y}_N + \text{tr}(\mathbf{C}_N^{-1}\mathbf{D}_N\mathbf{W}_{1N}), \\ \frac{1}{\tilde{\sigma}^2_N(\delta)}\Delta\widetilde{\mathbf{V}}'_N(\delta)\mathbf{C}_N^{-1}\mathbf{B}_{3N}(\lambda_3)\mathbf{W}_{2N}\Delta\mathbf{Y}_{N,-1} + \text{tr}(\mathbf{C}_N^{-1}\mathbf{D}_{N,-1}\mathbf{W}_{2N}), \\ \frac{1}{\tilde{\sigma}^2_N(\delta)}\Delta\widetilde{\mathbf{V}}'_N(\delta)(C_{T-1}^{-1}\otimes G_{3n}(\lambda_3))\Delta\widetilde{\mathbf{V}}_N(\delta) - (T-1)\text{tr}(G_{3n}(\lambda_3)), \\ \frac{1}{2\tilde{\sigma}^2_N(\delta)}\Delta\widetilde{\mathbf{V}}'_N(\delta)(C_{T-1}^{-1}\otimes \mathcal{Z}_{nj})\Delta\widetilde{\mathbf{V}}_N(\delta) - \frac{(T-1)}{2}Z'_{nj}1_n,\ j = 1, \ldots, k, \end{cases} \tag{4.9}$$

where $\delta = (\rho, \lambda')'$, $\Delta\widetilde{\mathbf{V}}_N(\delta) = \Delta\mathbf{V}_N(\tilde{\beta}_N(\delta), \delta)$, $\tilde{\beta}_N(\delta) = [\Delta\mathbb{X}'_N(\lambda_3)\Delta\mathbb{X}_N(\lambda_3)]^{-1}\Delta\mathbb{X}'_N(\lambda_3)\Delta\mathbb{Y}_N(\delta)$, $\tilde{\sigma}^2_N(\delta) = \frac{1}{N}\Delta\widetilde{\mathbf{V}}'_N(\delta)\mathbf{C}_N^{-1}\Delta\widetilde{\mathbf{V}}_N(\delta)$, $\Delta\mathbb{Y}_N(\delta) = \mathbf{C}_N^{-1/2}\mathbf{B}_{3N}(\lambda_3)[\mathbf{B}_{1N}(\lambda_1)\Delta\mathbf{Y}_N - \mathbf{B}_{2N}(\rho, \lambda_2)\Delta\mathbf{Y}_{N,-1}]$, $\Delta\mathbb{X}_N(\lambda_3) = \mathbf{C}_N^{-1/2}\mathbf{B}_{3N}(\lambda_3)\Delta\mathbf{X}_N$, and $\mathbf{C}_N^{1/2}$ is the symmetric square-root matrix of $\mathbf{C}_N$.

[14] Note that even when the errors are normal, this test does not have a simplified version as for SLR or FE-SPD model, as in this case the AQS function is not the true score function and the information matrix equality (IME) does not hold. The generalized IME cannot be applied as the true score function is unknown.

It is not difficult to see that $\text{plim}\frac{1}{N}S^{\circ c}_{\texttt{DSPD}}(\delta_0) = 0$, but $\text{E}[\frac{1}{N}S^{\circ c}_{\texttt{DSPD}}(\delta_0)] \neq 0$. Therefore, a direct use of $S^{\circ}_{\texttt{DSPD}}(\theta)$ to construct a test statistic would incur finite sample bias as it does not take into account the variability from the estimation of β and σ^2. To have a set of unbiased AQS functions for δ and α (at H_0), we work with $\tilde{\sigma}^2_N(\delta)S^{\circ c}_{\texttt{DSPD}}(\delta)$.

One can show that $\mathbf{C}_N^{-1/2}\Delta\widetilde{\mathbf{V}}_N(\delta_0) = \mathbf{M}_N\mathbf{C}_N^{-1/2}\Delta\mathbf{V}_N$ and $\tilde{\sigma}^2_N(\delta_0) = \frac{1}{N}\Delta\mathbf{V}'_N\mathbf{M}^{\circ}_N\Delta\mathbf{V}_N$, where $\mathbf{M}_N = I_N - \Delta\mathbb{X}_N(\Delta\mathbb{X}'_N\Delta\mathbb{X}_N)^{-1}\Delta\mathbb{X}'_N$, and $\mathbf{M}^{\circ}_N = \mathbf{C}_N^{-1/2}\mathbf{M}_N\mathbf{C}_N^{-1/2}$. By Lemma 3.1 of Yang (2018a): $\text{E}(\Delta\mathbf{Y}_{N,-1}\Delta\mathbf{V}'_N) = -\sigma_0^2\mathbf{D}_{N,-1}\mathbf{B}_{3N}^{-1}$ and $\text{E}(\Delta\mathbf{Y}_N\Delta\mathbf{V}'_N) = -\sigma_0^2\mathbf{D}_N\mathbf{B}_{3N}^{-1}$, one can easily show that $\text{E}[\tilde{\sigma}^2_N(\delta_0)S^{\circ c}_{\texttt{DSPD}}(\delta_0)] = \sigma_0^2(\mu^*_N + \frac{N-p}{N}\mu_N)$, where μ^*_N has the elements,

$$
\begin{aligned}
&\mu^*_{\rho} = -\text{tr}(\mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbf{D}_{N,-1}\mathbf{B}_{3N}^{-1}), && \mu^*_{\lambda_1} = -\text{tr}(\mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbf{W}_{1N}\mathbf{D}_N\mathbf{B}_{3N}^{-1}),\\
&\mu^*_{\lambda_2} = -\text{tr}(\mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbf{W}_{2N}\mathbf{D}_{N,-1}\mathbf{B}_{3N}^{-1}), && \mu^*_{\lambda_3} = \text{tr}(\mathbf{M}_N(I_{T-1}\otimes G_{3n})),\\
&\mu^*_{\alpha_j} = \tfrac{1}{2}\text{tr}(\mathbf{M}_N(I_{T-1}\otimes \mathcal{Z}_{nj})),\ j = 1,\ldots,k.
\end{aligned}
$$

and μ_N is the vector containing the second (non-stochastic) terms in (4.9).

Re-centering $\tilde{\sigma}^2_N(\delta_0)S^{\circ c}_{\texttt{DSPD}}(\delta_0)$ by subtracting it by $\frac{N}{N-p}\tilde{\sigma}^2_N(\delta_0)(\mu^*_N + \frac{N-p}{N}\mu_N)$ and simplifying, we obtain the further adjusted AQS functions:

$$
S^*_{\texttt{DSPD}}(\delta) = \begin{cases}
\Delta\widetilde{\mathbf{V}}'_N(\delta)\mathbf{C}_N^{-1}\mathbf{B}_{3N}(\lambda_3)\Delta\mathbf{Y}_{N,-1} - \mu^*_{\rho}(\delta)\tilde{\sigma}^{*2}_N(\delta),\\
\Delta\widetilde{\mathbf{V}}'_N(\delta)\mathbf{C}_N^{-1}\mathbf{B}_{3N}(\lambda_3)\mathbf{W}_{1N}\Delta\mathbf{Y}_N - \mu^*_{\lambda_1}(\delta)\tilde{\sigma}^{*2}_N(\delta),\\
\Delta\widetilde{\mathbf{V}}'_N(\delta)\mathbf{C}_N^{-1}\mathbf{B}_{3N}(\lambda_3)\mathbf{W}_{2N}\Delta\mathbf{Y}_{N,-1} - \mu^*_{\lambda_2}(\delta)\tilde{\sigma}^{*2}_N(\delta),\\
\Delta\widetilde{\mathbf{V}}'_N(\delta)\big[C_{T-1}^{-1}\otimes G_{3n}(\lambda_3)\big]\Delta\widetilde{\mathbf{V}}_N(\delta) - \mu^*_{\lambda_3}(\lambda_3)\tilde{\sigma}^{*2}_N(\delta),\\
\frac{1}{2}\Delta\widetilde{\mathbf{V}}'_N(\delta)\big(C_{T-1}^{-1}\otimes \mathcal{Z}_{nj}\big)\Delta\widetilde{\mathbf{V}}_N(\delta) - \mu^*_{\alpha_j}(\lambda_3)\tilde{\sigma}^{*2}_N(\delta),\ j = 1,\ldots,k,
\end{cases} \tag{4.10}
$$

where $\tilde{\sigma}^{*2}_N(\delta) = \frac{N}{N-p}\tilde{\sigma}^2_N(\delta)$. It is easy to show that $\text{E}[\tilde{\sigma}^{*2}_N(\delta_0)] = \sigma_0^2$ and $\text{E}[S^*_{\texttt{DSPD}}(\delta_0)] = 0$. Solving $\{S^*_{\texttt{DSPD},\delta}(\delta) = 0\}$ gives an adjusted AQS or M-estimator $\tilde{\delta}^*_{\texttt{M}}$ of δ, which has better finite sample properties. At the true δ_0, using (4.3), $S^*_{\texttt{DSPD}}(\delta_0)$ can be written as:

$$
S^*_{\texttt{DSPD}}(\delta_0) = \begin{cases}
\Delta\mathbf{V}'_N\mathbf{\Psi}^*_1\Delta\mathbf{Y}_{N1} + \Pi^{*\prime}_1\Delta\mathbf{V}_N + \Delta\mathbf{V}'_N\mathbf{\Phi}^*_1\Delta\mathbf{V}_N,\\
\Delta\mathbf{V}'_N\mathbf{\Psi}^*_2\Delta\mathbf{Y}_{N1} + \Pi^{*\prime}_2\Delta\mathbf{V}_N + \Delta\mathbf{V}'_N\mathbf{\Phi}^*_2\Delta\mathbf{V}_N,\\
\Delta\mathbf{V}'_N\mathbf{\Psi}^*_3\Delta\mathbf{Y}_{N1} + \Pi^{*\prime}_3\Delta\mathbf{V}_N + \Delta\mathbf{V}'_N\mathbf{\Phi}^*_3\Delta\mathbf{V}_N,\\
\Delta\mathbf{V}'_N\mathbf{\Phi}^*_4\Delta\mathbf{V}_N,\\
\Delta\mathbf{V}'_N\mathbf{\Phi}^*_{4+j}\Delta\mathbf{V}_N,\ j = 1,\ldots,k,
\end{cases} \tag{4.11}
$$

where $\mathbf{\Phi}^*_{4+j} = \frac{1}{2}\mathbf{M}^{\circ}_N(C_{T-1}\otimes\mathcal{Z}_{nj})\mathbf{M}^{\circ}_N - \frac{1}{N-p}\mu^*_{\alpha_j}\mathbf{M}^{\circ}_N,\ j = 1,\ldots,k,$

$$
\begin{aligned}
&\Pi^*_1 = \mathbf{M}^{\circ}_N\mathbf{B}_{3N}\boldsymbol{\eta}_{-1}, && \mathbf{\Psi}^*_1 = \mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbb{R}_{-1}, && \mathbf{\Phi}^*_1 = \mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbb{S}_{-1} - \tfrac{1}{N-p}\mu^*_{\rho}\mathbf{M}^{\circ}_N,\\
&\Pi^*_2 = \mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbf{W}_{1N}\boldsymbol{\eta}, && \mathbf{\Psi}^*_2 = \mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbf{W}_{1N}\mathbb{R}, && \mathbf{\Phi}^*_2 = \mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbf{W}_{1N}\mathbb{S} - \tfrac{1}{N-p}\mu^*_{\lambda_1}\mathbf{M}^{\circ}_N,\\
&\Pi^*_3 = \mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbf{W}_{2N}\boldsymbol{\eta}_{-1}, && \mathbf{\Psi}^*_3 = \mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbf{W}_{2N}\mathbb{R}_{-1}, && \mathbf{\Phi}^*_3 = \mathbf{M}^{\circ}_N\mathbf{B}_{3N}\mathbf{W}_{2N}\mathbb{S}_{-1} - \tfrac{1}{N-p}\mu^*_{\lambda_2}\mathbf{M}^{\circ}_N,\\
&\mathbf{\Phi}^*_4 = \mathbf{M}^{\circ}_N(C_{T-1}\otimes G_{3n})\mathbf{M}^{\circ}_N - \tfrac{1}{N-p}\mu^*_{\lambda_3}\mathbf{M}^{\circ}_N.
\end{aligned}
$$

The AQS function (4.11) has a similar structure as (4.4), containing the same three types of terms so that (4.5)-(4.7) can be applied. Using these newly defined quantities, an MD representation can be developed for $S^*_{\texttt{DSPD}}(\delta_0)$, i.e., $S^*_{\texttt{DSPD}}(\delta_0) = \sum_{i=1}^n \mathbf{g}^*_{ni}$, which gives $\mathbf{\Omega}^*_N = \text{Var}[S^*_{\texttt{DSPD}}(\delta_0)] = \sum_{i=1}^n \text{E}(\mathbf{g}^*_{ni}\mathbf{g}^{*\prime}_{ni})$ and hence a consistent OPMD estimate of $\mathbf{\Omega}^*_N$ as $\widetilde{\mathbf{\Omega}}^*_N = \sum_{i=1}^n \tilde{\mathbf{g}}^*_{ni}\tilde{\mathbf{g}}^{*\prime}_{ni}$. Partition $S^*_{\texttt{DSPD}}(\delta)$, $\mathbf{g}^*_{ni}$ and $\mathbf{\Omega}^*_N$ according to δ and α. Let $\mathbf{\Sigma}^*_{N,\alpha\delta} = -\text{E}[\frac{\partial}{\partial\delta'} S^*_{\texttt{DSPD},\alpha}(\delta_0)]$, and $\mathbf{\Sigma}^*_{N,\delta\delta} = -\text{E}[\frac{\partial}{\partial\delta'} S^*_{\texttt{DSPD},\delta}(\delta_0)]$. A potentially improved test statistic can be constructed on the basis of $S^*_{\texttt{DSPD}}(\delta_0)$ in an identical manner as for $T^{\text{r}}_{\texttt{DSPD}}$:

$$T^{\text{r}*}_{\texttt{DSPD}} = \big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}^{*\prime}_{ni,\alpha}\big)\big[\textstyle\sum_{i=1}^n (\tilde{\mathbf{g}}^*_{ni,\alpha} - \widetilde{\mathbf{\Gamma}}^*_N \tilde{\mathbf{g}}^*_{ni,\delta})(\tilde{\mathbf{g}}^*_{ni,\alpha} - \widetilde{\mathbf{\Gamma}}^*_N \tilde{\mathbf{g}}^*_{ni,\delta})'\big]^{-1}\big(\textstyle\sum_{j=1}^n \tilde{\mathbf{g}}^*_{ni,\alpha}\big), \qquad (4.12)$$

where $\tilde{\mathbf{g}}^*_{ni,\alpha}$ and $\tilde{\mathbf{g}}^*_{ni,\delta}$ are the estimates of $\mathbf{g}^*_{ni,\alpha}$ and $\mathbf{g}^*_{ni,\delta}$ at $\tilde{\delta}^*_N$, $\widetilde{\mathbf{\Gamma}}^*_N = \widetilde{\mathbf{\Sigma}}^*_{N,\alpha\delta}\widetilde{\mathbf{\Sigma}}^{*-1}_{N,\delta\delta}$, $\widetilde{\mathbf{\Sigma}}^*_{N,\alpha\delta} = -\frac{\partial}{\partial\delta'} S^*_{\texttt{DSPD},\alpha}(\tilde{\delta}^*_{\texttt{M}})$, and $\widetilde{\mathbf{\Sigma}}^*_{N,\delta\delta} = -\frac{\partial}{\partial\delta'} S^*_{\texttt{DSPD},\delta}(\tilde{\delta}^*_{\texttt{M}})$, where $\tilde{\delta}^*_{\texttt{M}}$ is the improved M-estimator of δ defined above. These derivatives can be obtained from (4.10), taking use of the expression $\Delta\widetilde{\mathbf{V}}_N(\delta) = \mathbf{C}_N^{1/2}\mathbf{M}_N(\lambda_3)\Delta\mathbb{Y}_N(\delta)$. See Appendix D for details.

Theorem 4.2. *Under the assumptions of Theorem 4.1, we have under* H_0, $T^{r*}_{\texttt{DSPD}} \xrightarrow{D} \chi^2_k$.

5. Monte Carlo Study

Extensive Monte Carlo experiments are performed for assessing the finite sample performance of the four tests proposed in Section 2 for the SLR model, the six tests introduced in Section 3 for the FE-SPD model, and the two tests presented in Section 4 for the FE-DSPD model. An important purpose is to solicit accurate and reliable tests based on the Monte Carlo results, and to make recommendations for practitioners.

5.1. General Settings

Cross-Sectional Case. We use the SLR model (2.1), where the matrix X_n contains a constant (ι_n) and one regressor (x_n). Throughout the experiment the parameters are set at $\beta_0 = 5$, $\beta_1 = 1$, λ_1, $\lambda_2 = 0.2$, 0.8, and $n = 50, 100$, 200 and 500. For the spatial weight matrices, we assume that $W_{1n} = W_{2n} = W_n$. We have taken the spatial matrix W_n proposed by Kelejian and Prucha (1999), which is labelled "J ahead and J behind" with the non-zero elements being $1/2J$. Clearly, as J increases, the number of non-zero elements in the spatial weight matrix increases, which is in turn increases the 'degree' of the spatial dependence. Moreover, following Baltagi and Yang (2013a), we have also considered three other schemes for generating the spatial weights matrices: (i) Rook contiguity, (ii) Queen contiguity and (iii) Group interactions. In the last one, the degree of spatial dependence grows with the sample size, which is achieved by relating the number of groups k to the sample size n, e.g., $k = n^{0.5}$, see Lee (2004). Two Data Generating Processes (DGP) are considered to generate the elements $\{x_i\}$ of the regressors x_n. The first one (DGP1) assumes that $\{x_i\}$ are *iid* $N(0,1)$, whereas the second one (DGP2) considers that there might be systematic differences in $\{x_i\}$ across the different 'sets' of spatial units, see Baltagi and Yang (2013a) and Lee (2004). In this case, the ith value in the jth group, $\{x_{ij}\}$ of x_n are generated according to $\{x_{ij}\} = (z_j + \epsilon_{ij})/\sqrt{2}$ where $\{z_j\} \sim$ *iid* $N(0,1)$, $\{\epsilon_{ij}\}$ are *iid* $N(0,1)$, and z_j and ϵ_{ij} are independent. This second scheme gives non-*iid* $\{x_i\}$ values in contrast to the first one, or different group means in terms of group interaction, see Lee (2004). The heteroskedasticity is generated according to $\sigma^2_{v_{ni}} = \sigma^2 \exp(\alpha z_{ni})$, where z_{ni} is taken to be x_{ni}, σ is set to 1, and $\alpha = 0, 1$. If $\alpha = 0$, the disturbances are homoskedastic. For the DGP of disturbances, we assume that $v_{ni} = \sigma^2_{v_{ni}} e_i$, where $\{e_i\}$ are generated from either $N(0,1)$, or a chi-square distribution with 3 degrees of freedom, standardized to have zero mean and unit variance.

Static Panel Case. We use the FE-SPD model (3.1), which includes a single time-varying regressor X_{nt}. The fixed effects are generated by setting $\mu_n = \frac{1}{T}\sum_{t=1}^{T} X_{nt} + \omega_n$ where $\omega_n \sim N(0, I_n)$. Two DGPs are also considered for generating the regressor' values. In DGP1, we have $x_{it} = z_{it} + 0.1t$, where $\{z_{it}\}$ are *iid* $N(0,1)$. Thus the regressor includes a time trend $0.1t$. In DGP2, we first generate X_{nt} for each t according to the DGP2 for the SLR model and then add a time trend $0.1t$ on each X_{nt}, $t = 1, \ldots, T$. Four individual dimensions are considered $n = 50$, 100, 200 and 500 combined with the time dimension $T = 5$. Throughout the experiment the parameters are set at $\beta = 1$, λ_1, $\lambda_2 = 0.2$, 0.8. The spatial matrices are those that have been defined for the SLR model. The heteroskedasticity is generated according to $\sigma^2_{v_{ni}} = \sigma^2 \exp(\alpha z_{ni})$, where $z_{ni} = \frac{1}{T}\sum_{t=1}^{T} x_{it}$, $\sigma = 1$, and $\alpha = 0, 0.5$. If $\alpha = 0$, the disturbances are homoskedastic. For the DGP of the disturbances, we assume that $v_{n,it} = \sigma^2_{v_{ni}} e_{it}$, where $\{e_{it}\}$ are generated from either $N(0,1)$ or a chi-square distribution with 3 degrees of freedom, standardized to have zero mean and unit variance.

Dynamic Panel Case. As discussed in Sections 1 and 4, this is the most important case (among the three studied in the paper) used to demonstrate the usefulness of the introduced AQS-OPMD methodology in constructing tests of homoskedasticity in 'non-standard' situations. We now use the Monte Carlo tool to assess how the introduced tests perform in finite samples. We use the FE-DSPD model (4.1), which contains a single time-varying regressor X_{nt} generated in the same way as for the FE-SPD model, i.e. by DGP1 and DGP2. The same individual and time dimensions are also retained, with the exception of $n = 500$ which is replaced by $n = 400$. This substitution simply aims to save time in simulations. Throughout the experiment the parameters are set at $\beta = 1$, $\rho = 0.3$, and $(\lambda_1, \lambda_2, \lambda_3) = (0.2, 0.2, 0.2)$ or $(-0.2, -0.2, -0.2)$. The spatial matrices are those that have been defined for the FE-SPD model as well as the heteroskedasticity disturbance processes, where the heteroskedasticity parameter α is set at 0 (homoskedasticity) for size simulation and 0.5 for power simulation.

5.2. Monte Carlo Results

In all the experiments, the regressors are treated as fixed. Each set of results, corresponding to a combination of the value of n, the values of (λ_1, λ_2) for SLR and FE-SPD models and also (ρ, λ_3) for the FE-DSPD model, a DGP, a set of spatial weight matrices and an error distribution, is based on $5,000$ Monte Carlo replications. Three nominal sizes are considered: 10%, 5% and 1%. Empirical size and size adjusted power of the tests are recorded. Due to space constraint, only partial results are reported, corresponding to queen and group interactions spatial layouts. Other results, corresponding to DGP2 and $(\lambda_1, \lambda_2) = (0.8, 0.2)$ for SLR and FE-SPD models, and circular world and rook contiguity spatial layouts for all three models, are available upon request from the authors.

Cross-Sectional Case. Tables 1 and 2 report partial results on, respectively, the empirical sizes and the empirical size adjusted powers of the four tests: $T_{\texttt{SLR}}$, $T^{\text{r}}_{\texttt{SLR}}$, $T^{*}_{\texttt{SLR}}$ and $T^{\text{r}*}_{\texttt{SLR}}$, introduced in Section 2 for the SLR model. From the results (reported and unreported), the following general observations are in order:

(*i*) Among the four tests, the AQS test $T^{\text{r}*}_{\texttt{SLR}}$ performs the best in the sense that its empirical sizes are in general quite close to the corresponding nominal levels. The score test $T_{\texttt{SLR}}$ performs the worst, much worse than the other three in terms of size;

(*ii*) Non-normality can have a big impact on the finite sample performance of the tests – size distortion can be much bigger when the errors are non-normal than when they are normal, except for the AQS test $T^{\text{r}*}_{\texttt{SLR}}$ where the size distortions are at an 'acceptable' level even when $n = 50$;

(*iii*) When the errors are normal, the size converges to its nominal level, for all tests considered, as the sample size n increases. When the errors are non-normal, the two robust tests converge as expected. For the two non-robust tests, the score test $T_{\mathtt{SLR}}$ still has a large size distortion even when the sample size is 500, but the AS test $T^*_{\mathtt{SLR}}$ has size quite close to its nominal level when n is large enough, showing that it is fairly robust;

(*iv*) Neither the values of spatial parameters nor the spatial weight matrices have a significant effect on the finite sample performance of the tests. One exception is that under normality and when n is not large, the last three tests can be slightly under-sized;

(*v*) As n increases, all four tests have empirical **size adjusted powers** converging to 100%. As expected, the two finite sample adjusted tests (AS and AQS) have lower powers than the other two (score and QS), but as n increases, their powers quickly catch up.

Comparing the quasi score test $T^r_{\mathtt{SLR}}$ with the score test $T_{\mathtt{SLR}}$ (see Section 2.2), we see that the simple changes on $T_{\mathtt{SLR}}$ not only offer robustness against non-normality but also lead to huge improvements in its finite sample performance. Comparing the adjusted score test $T^*_{\mathtt{SLR}}$ with the score test $T_{\mathtt{SLR}}$, we see that some simple adjustments on the concentrated scores can lead to huge improvements in the finite sample performance of the test. Thus, a combination of the idea leading to the AS test and the idea leading to the QS test, we obtain an AQS test that not only is robust against non-normality but also has the best finite sample properties. In light of the overall performance, the AQS test $T^{\mathrm{r}*}_{\mathtt{SLR}}$ is recommended for practical applications.

Static Panel Case. Tables 3 and 4 present partial results on, respectively, the empirical sizes and the empirical size adjusted powers for the six tests introduced in Section 3 pertaining to the FE-SPD model: $T_{\mathtt{SPD}}$, $T^{\mathrm{r}}_{\mathtt{SPD}}$, $T^{\mathrm{rr}}_{\mathtt{SPD}}$, $T^*_{\mathtt{SPD}}$, $T^{\mathrm{r}*}_{\mathtt{SPD}}$ and $T^{\mathrm{rr}*}_{\mathtt{SPD}}$.

The results (reported and unreported) show similar patterns for the FE-SPD model as for the SLR model. In particular, the score test $T_{\mathtt{SPD}}$ can have large size distortions when n is small and the errors are non-normal, irrespective of the values of the spatial parameters, the spatial weight matrix structures, and the way the regressor is generated. Similar patterns are observed for the tests $T^{\mathrm{r}}_{\mathtt{SPD}}$, $T^*_{\mathtt{SPD}}$, and $T^{\mathrm{r}*}_{\mathtt{SPD}}$, though the size-distortions are on a smaller scale when compared with the score test. The sizes of these four tests do not seem to converge to the nominal levels as the large size distortions remain even when $n = 500$ with $T = 5$.

In contrast, the two fully robust tests $T^{\mathrm{rr}}_{\mathtt{SPD}}$ and $T^{\mathrm{rr}*}_{\mathtt{SPD}}$ in general offer a great reduction in size distortion and their power converges to 100% as n increases. The empirical sizes of these two tests converge to their nominal levels as $n \to \infty$ where T can go large with n or stay fixed. Hence the two fully robust tests are both recommended for practical applications.

Dynamic Panel Case. Tables 5 and 6 report partial results on the empirical sizes and the empirical size adjusted powers for the two tests introduced in Section 4 pertaining to the FE-DSPD model: $T^{\text{r}}_{\texttt{DSPD}}$ and $T^{\text{r}*}_{\texttt{DSPD}}$, under DGP1 and DGP2, respectively.

By construction, both tests are robust against non-normality. Therefore, it is important to assess and to compare their finite sample performance. The results (reported and unreported) show that both tests perform well in finite sample in that the empirical sizes are all quite close to their nominal levels even when n is as small as 50. The (surprisingly) good performance of $T^{\text{r}}_{\texttt{DSPD}}$ is perhaps due to the fact that the AQS function used to construct $T^{\text{r}}_{\texttt{DSPD}}$ has already gone through some major adjustments, or the fact that $\dim(\beta)$ is low so that the effect of its estimation is small in this situation. The results (reported and unreported) further show that $T^{\text{r}*}_{\texttt{DSPD}}$ is slightly less powerful than $T^{\text{r}}_{\texttt{DSPD}}$ when n is not large, but both converge to 100% very fast. Both tests are very powerful. In fact, for $\alpha = 0.5$ all the empirical powers are quite close to 100%, showing that FE-DSPD model is very sensitive to departures from homoskedasticity. Thus, both tests can be used in practical applications, with $T^{\text{r}*}_{\texttt{DSPD}}$ being more preferred when $\dim(\beta)$ is big.

6. Conclusion

We introduce an *Adjusted Quasi-Score* (AQS) method for constructing diagnostic tests for homoskedasticity in spatial econometric models, by first adjusting the score-type function from a given model to give a set of AQS functions that are unbiased and consistent, and then developing an *Outer-Product-of-Martingale-Difference* (OPMD) estimate of its variance. We use a spatial cross-section model to demonstrate that in standard problems where a genuine (quasi) score vector is available, the AQS method leads to finite sample improved tests over the usual methods by adjusting the concentrated (quasi) score to remove the effect of estimating *nuisance* parameters. We then consider a "not-so-standard problem", the spatial panel data model with fixed effects where the transformed errors are not totally independent, to further demonstrate its ability to yield finite sample improved tests. Finally, we focus on a "non-standard problem", the spatial dynamic panel data model, to demonstrate that the AQS-OPMD method is able to provide a feasible solution to non-standard problems where the standard methods fail. Asymptotic properties of the tests developed for these three models are formally studied, and Monte Carlo results show that our testing procedures perform well in finite samples, especially for the robust versions of the tests. Based on our theoretical arguments and empirical findings, we recommend the following:

- for SLR models, the AQS test $T^{r*}_{\texttt{SLR}}$, given in (2.19), should be used;
- for SPD models, the two fully robust AQS tests $T^{rr}_{\texttt{SPD}}$ and $T^{rr*}_{\texttt{SPD}}$, given respectively in (3.14) and (3.19), can both be used;
- for DSPD models, the AQS tests $T^{r}_{\texttt{DSPD}}$ and $T^{r*}_{\texttt{DSPD}}$, given respectively in (4.8) and (4.12), can both be used.

The method is seen to be quite general. It can be easily applied to many other standard problems to take into account the time-wise fixed effects and heteroskedasticity, spatial Durbin effects, higher-order spatial lags and spatial errors. It has potential to provide feasible solutions to testing problems in many other non-standard problems such as spatial panels with temporal heterogeneity (time-specific fixed effects, time-varying spatial weights, time-varying regression and spatial coefficients, etc.), unbalanced spatial panels, spatial panels with interactive fixed effects, etc. Moreover, the tests can be repeatedly run with different choices of the heteroskedasticity variables, to identify the 'source' of heteroskedasticity: the heteroskedasticity variables with which the test is rejected. In this case, one may proceed with a heteroskedastic model by 'specifying' a form for the unknown function $h(\cdot)$, e.g., the popular exponential form, or non-parametrically estimating it.

Appendix. Supplementary data

A supplementary material, containing Appendices A, B, C, and D that the paper refers to, is available from the journal's website.

Acknowledgements

We would like to thank Gianfranco Piras, James LeSage, and the participants of the XII World Conference of the Spatial Econometrics Association, WVPU Vienna, June 2018, for their valuable suggestions and helpful comments. We are grateful to Editor Serena Ng for her encouragements in revising the paper, and for her helpful and constructive comments that have led to a much enhanced version of the paper. Helpful comments from three anonymous referees and two associate editors are also gratefully acknowledged. Zhenlin Yang gratefully acknowledges the research support from Singapore Management University under Grant C244/MSS16E003.

References

[1] Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht.

[2] Badinger, H., Egger, P., 2011. Estimation of higher-order spatial autoregressive cross-section models with heteroskedastic disturbances. *Papers in Regional Science* 90, 213-235.

[3] Badinger, H., Egger, P., 2015. Fixed effects and random effects estimation of higher-order spatial autoregressive models with spatial autoregressive and heteroskedastic disturbances. *Spatial Economic Analysis* 10:1, 11-35.

[4] Bai, J., Ng, S., 2013. Principal components estimation and identification of static factors. *Journal of econometrics* 176, 18-29.

[5] Baltagi, B. H., Yang, Z. L., 2013a. Standardized LM tests of spatial error dependence in linear or panel regressions. *The Econometrics Journal* 16, 103-134.

[6] Baltagi, B.H., Yang, Z. L., 2013b. Heteroskedasticity and non-normality robust LM tests for spatial dependence. *Regional Science and Urban Economics* 43, 725-739.

[7] Breusch, T.S., Pagan, A. R., 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica* 47, 1287-1293.

[8] Davidson, J., 1994. *Stochastic Limit Theory*. Oxford University Press, Oxford.

[9] Debarsy, N., Jin, F., Lee, L.-F, 2015. Large sample properties of the matrix exponential spatial specification with an application to FDI. *Journal of Econometrics* 188, 1-21.

[10] Elhorst, J. P., 2014. *Spatial Econometrics from Cross-Sectional Data to Spatial Panels*. Springer.

[11] Hsiao, C., 2014. *Analysis of Panel Data*. Third Edition, Cambridge University Press, New York.

[12] Hsiao, C., Pesaran, M. H., Tahmiscioglu, A. K., 2002. Maximum likelihood estimation of fixed effects dynamic panel data models covering short panels. *Journal of Econometrics* 109, 107-150.

[13] Jin, F., Lee, L.-F., 2013. Cox-type tests for competing spatial autoregressive models with spatial autoregressive disturbances. *Regional Science and Urban Economics* 43, 590-616.

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

[15] Kelejian, H. H., Prucha, I. R., 2001. On the asymptotic distribution of the Moran *I* test statistic with applications. *Journal of Econometrics* 104, 219-257.

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

[17] Kuersteiner, G. M., Prucha, I. R., 2020. Dynamic panel data models: networks, common shocks, and sequential exogeneity. *Econometrica*, Forthcoming.

[18] Lee, L.-F., 2002. Consistency and efficiency of least squares estimation for mixed regressive spatial autoregressive models. *Econometric Theory* 18, 252-277.

[19] Lee, L.-F., 2004. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* 72, 1899-1925.

[20] Lee, L.-F., 2007. Identification and estimation of econometric models with group interactions, contextual factors and fixed effects. *Journal of Econometrics* 140, 333-374.

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

[22] Lee, L.-F., Yu, J., 2016. Identification of spatial Durbin panel models. *Journal of Applied Econometrics* 31, 133-162.

[23] Lin, X., Lee, L.-F., 2010. GMM estimation of spatial autoregressive models with unknown heteroskedasticity. *Journal of Econometrics* 157, 34-52.

[24] Liu, S. F., Yang, Z. L., 2015a. Asymptotic distribution and finite-sample bias correction of QML estimators for spatial error dependence model. *Econometrics* 3, 376-411.

[25] Liu, S. F., Yang, Z. L., 2015b. Modified QML Estimation of Spatial Autoregressive Models with Unknown Heteroskedasticity and Nonnormality. *Regional Science and Urban Economics*. 52, 50-70.

[26] Moscone, F., Tosetti, E., 2011. GMM estimation of spatial panels with fixed effects and unknown heteroskedasticity. *Regional Science and Urban Economics* 41 487-497.

[27] Neyman, J., Scott, E. L., 1948. Consistent estimates based on partially consistent observations. *Econometrica* 16, 1-32.

[28] Su, L., Yang, Z. L., 2015. QML estimation of dynamic panel data models with spatial errors. *Journal of Econometrics* 185, 230-258.

[29] Yang, Z. L., 2015a. LM tests of spatial dependence based on bootstrap critical values. *Journal of Econometrics* 185, 33-59.

[30] Yang, Z. L., 2015b. A general method for third-order bias and variance correction on a nonlinear estimator. *Journal of Econometrics* 186, 178-200.

[31] Yang, Z. L., 2018a. Unified M-estimation of fixed-effects spatial dynamic models with short panels. *Journal of Econometrics* 205, 423-447.

[32] Yang, Z. L., 2018b. Bootstrap LM tests for higher order spatial effects in spatial linear regression models. *Empirical Economics* 55, 35-68.

[33] Yang, Z. L., J. Yu and S.F. Liu, 2016. Bias correction and refined inferences for fixed effects spatial panel data models. *Regional Science and Urban Economics* 61, 52-72.

[34] Yu, J., de Jong, R., Lee, L. F., 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.

Table 1 – Empirical Size of the Tests for the SLR Model, **DGP1**

			Queen contiguity						Group interaction					
			Normal disturbances			Non-normal disturbances			Normal disturbances			Non-normal disturbances		
(λ_1, λ_2)	n	Tests	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
(0.2,0.2)	50	T_{SLR}	22.0	13.66	4.86	40.78	31.26	17.00	22.72	14.86	4.62	42.08	32.48	17.94
	50	T_{SLR}^{r}	9.70	5.42	1.52	15.08	9.22	3.06	11.74	5.90	1.20	18.04	11.22	3.98
	50	T_{SLR}^{*}	12.24	5.92	1.38	19.44	10.36	1.66	13.38	7.70	1.58	20.44	11.74	3.18
	50	T_{SLR}^{r*}	9.72	4.46	0.86	13.98	6.68	0.82	11.06	5.52	1.00	14.58	7.58	1.44
	100	T_{SLR}	17.72	10.62	3.20	37.86	29.00	15.42	17.30	9.98	2.88	38.12	29.52	15.84
	100	T_{SLR}^{r}	8.94	4.20	0.86	13.76	7.62	2.36	9.12	4.54	0.84	14.92	9.28	3.30
	100	T_{SLR}^{*}	12.00	5.90	1.00	15.86	7.56	1.62	11.64	5.82	1.08	16.38	9.36	2.18
	100	T_{SLR}^{r*}	10.68	4.98	0.66	12.28	5.56	1.16	9.90	4.98	0.76	12.68	7.06	1.20
	200	T_{SLR}	13.30	7.18	1.82	33.80	25.52	14.20	13.86	7.82	2.26	33.30	24.78	12.60
	200	T_{SLR}^{r}	8.52	4.36	0.98	13.28	7.62	2.26	10.74	5.46	1.20	13.76	7.42	2.18
	200	T_{SLR}^{*}	10.26	5.00	0.88	13.54	7.00	1.18	11.00	5.34	1.06	13.02	6.72	1.26
	200	T_{SLR}^{r*}	9.72	4.52	0.74	11.48	5.58	1.02	10.22	4.98	0.90	10.80	4.90	0.74
	500	T_{SLR}	11.98	6.08	1.62	29.42	20.70	9.34	12.26	6.54	1.40	30.54	21.38	10.12
	500	T_{SLR}^{r}	10.08	4.82	1.08	11.82	6.18	1.32	10.08	5.12	0.98	12.20	6.66	1.48
	500	T_{SLR}^{*}	10.42	4.82	1.02	11.24	5.74	1.02	10.78	5.36	0.98	12.06	6.02	1.06
	500	T_{SLR}^{r*}	10.26	4.68	1.00	10.28	5.22	0.88	10.42	5.18	0.98	11.08	5.34	0.80
(0.2,0.8)	50	T_{SLR}	21.80	13.44	5.06	41.80	32.76	18.46	22.82	15.04	5.16	40.92	32.28	17.92
	50	T_{SLR}^{r}	12.00	6.44	1.70	21.10	14.18	4.86	14.04	8.00	1.68	21.78	14.36	5.08
	50	T_{SLR}^{*}	12.38	5.74	1.30	19.90	9.84	1.48	14.58	7.32	1.10	16.40	9.40	2.40
	50	T_{SLR}^{r*}	9.32	4.60	0.84	13.68	6.04	0.68	9.32	4.00	0.42	10.38	4.88	1.10
	100	T_{SLR}	16.44	10.02	2.98	36.58	28.88	15.56	17.26	10.10	2.82	38.02	28.42	16.18
	100	T_{SLR}^{r}	11.28	6.10	1.42	16.26	9.82	3.10	16.13	6.72	1.54	18.70	11.70	3.92
	100	T_{SLR}^{*}	11.52	5.98	0.92	15.86	7.84	1.30	10.46	4.90	0.86	14.72	8.02	1.80
	100	T_{SLR}^{r*}	10.34	4.82	0.56	11.82	5.08	0.80	8.84	3.64	0.54	10.74	4.78	0.82
	200	T_{SLR}	13.72	7.88	2.04	34.10	25.42	13.30	14.04	7.64	2.16	32.72	23.98	12.20
	200	T_{SLR}^{r}	10.74	5.68	1.04	15.20	8.38	2.36	11.84	6.00	1.38	14.04	8.32	2.40
	200	T_{SLR}^{*}	10.56	5.28	0.76	13.26	6.72	1.30	11.12	5.44	0.86	13.34	6.92	1.32
	200	T_{SLR}^{r*}	9.76	5.10	0.58	10.98	5.30	0.94	9.88	4.62	0.68	10.24	4.86	0.78
	500	T_{SLR}	11.72	6.24	1.42	27.84	19.28	8.66	11.32	6.18	1.24	28.54	20.00	9.22
	500	T_{SLR}^{r}	10.22	5.12	1.12	11.56	6.02	1.18	10.38	5.44	1.04	11.68	6.20	1.24
	500	T_{SLR}^{*}	10.36	5.16	1.00	11.44	5.62	0.88	10.14	4.96	0.78	11.48	5.72	0.82
	500	T_{SLR}^{r*}	10.00	5.02	0.96	9.92	4.76	0.64	9.38	4.28	0.64	9.76	4.50	0.48
(0.8,0.2)	50	T_{SLR}	21.84	13.56	4.86	42.32	32.20	18.32	21.94	13.72	4.42	40.72	31.36	17.12
	50	T_{SLR}^{r}	10.68	5.84	1.66	19.58	12.28	4.00	10.24	5.30	1.08	18.38	11.58	3.92
	50	T_{SLR}^{*}	12.28	6.02	1.20	18.40	10.60	2.54	11.24	5.36	1.14	17.80	9.32	1.74
	50	T_{SLR}^{r*}	9.52	3.98	0.76	13.66	7.22	1.22	8.18	3.46	0.50	12.02	4.94	0.62
	100	T_{SLR}	16.82	10.02	2.72	38.00	29.46	16.32	16.62	9.62	2.74	38.30	29.32	15.92
	100	T_{SLR}^{r}	8.52	4.18	0.58	15.38	9.08	2.94	9.62	4.46	1.18	15.80	9.36	3.02
	100	T_{SLR}^{*}	11.24	5.26	0.82	15.56	7.66	1.46	11.46	4.88	0.80	15.90	8.32	1.50
	100	T_{SLR}^{r*}	10.00	4.44	0.50	12.10	5.56	0.96	9.80	4.08	0.60	12.26	5.76	0.70
	200	T_{SLR}	14.48	8.16	2.06	32.72	24.00	12.36	13.44	7.30	1.54	35.80	26.58	13.78
	200	T_{SLR}^{r}	9.76	5.24	0.98	12.08	6.74	1.72	10.14	4.86	1.04	15.02	8.08	2.28
	200	T_{SLR}^{*}	11.44	5.54	0.92	12.48	6.20	1.22	10.58	4.78	0.72	14.06	6.82	1.18
	200	T_{SLR}^{r*}	10.82	5.10	0.76	10.64	5.38	0.88	9.76	4.08	0.62	11.00	5.46	0.88
	500	T_{SLR}	11.44	5.82	1.48	29.08	21.00	10.50	11.34	5.80	1.48	29.86	20.72	9.62
	500	T_{SLR}^{r}	9.32	4.62	1.20	11.74	6.24	1.34	9.74	4.76	1.18	11.48	6.16	1.82
	500	T_{SLR}^{*}	10.16	5.02	1.00	11.92	5.86	1.00	9.90	4.72	1.10	11.54	5.56	1.32
	500	T_{SLR}^{r*}	9.86	4.90	0.96	10.84	5.04	0.86	9.76	4.54	0.96	10.26	4.96	1.06
(0.8,0.8)	50	T_{SLR}	20.66	13.02	4.60	41.90	32.54	18.24	22.58	14.58	4.74	43.44	34.60	19.48
	50	T_{SLR}^{r}	12.76	7.42	1.62	21.76	14.08	4.80	13.98	7.76	1.84	23.44	15.26	6.02
	50	T_{SLR}^{*}	12.14	5.48	0.98	18.12	9.06	1.86	14.06	7.60	1.28	19.56	10.94	3.44
	50	T_{SLR}^{r*}	9.18	3.98	0.50	11.66	4.98	0.92	9.26	4.32	0.50	9.98	4.66	0.78
	100	T_{SLR}	16.60	9.64	2.44	37.02	28.30	15.46	16.86	9.74	3.18	40.16	31.40	17.14
	100	T_{SLR}^{r}	11.86	6.24	0.96	17.36	10.52	3.24	12.80	6.78	1.78	19.16	11.74	3.88
	100	T_{SLR}^{*}	11.48	5.06	0.68	15.74	8.20	1.32	10.32	4.92	0.86	15.54	8.54	2.40
	100	T_{SLR}^{r*}	10.00	4.12	0.36	11.74	5.66	0.62	7.64	3.28	0.54	10.62	5.20	1.40
	200	T_{SLR}	14.18	7.58	1.58	34.66	25.46	13.00	14.14	8.28	2.02	34.94	25.48	13.28
	200	T_{SLR}^{r}	11.50	5.82	1.04	14.82	8.80	2.62	11.62	6.48	1.32	15.62	8.98	2.54
	200	T_{SLR}^{*}	10.80	4.96	0.64	13.38	6.40	1.40	11.38	5.70	0.78	15.12	7.86	1.64
	200	T_{SLR}^{r*}	9.96	4.52	0.54	10.86	5.06	0.90	10.38	4.74	0.64	11.68	5.58	0.78
	500	T_{SLR}	11.96	6.40	1.60	28.22	19.72	9.28	16.70	10.32	3.10	38.60	29.80	16.08
	500	T_{SLR}^{r}	10.56	5.80	1.34	11.84	6.14	1.36	12.64	6.86	1.84	18.40	11.00	3.32
	500	T_{SLR}^{*}	10.38	5.50	1.16	11.48	5.60	0.84	11.92	5.80	1.18	16.60	8.36	1.40
	500	T_{SLR}^{r*}	10.10	5.34	1.10	10.02	4.40	0.70	8.72	3.84	0.52	11.18	4.70	0.50

Table 2 – Empirical Size Adjusted Powers of the Tests for the SLR Model, **DGP1 - Disturbances Heteroskedastic $\alpha = 1.0$**

			Queen contiguity						Group interaction					
			Normal disturbances			Non-normal disturbances			Normal disturbances			Non-normal disturbances		
(λ_1,λ_2)	n	Tests	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
(0.2,0.2)	50	T_{SLR}	99.34	98.74	94.32	86.86	75.72	44.78	95.02	91.54	79.56	73.14	59.86	32.40
	50	T^{r}_{SLR}	91.58	89.00	81.28	81.06	70.98	46.20	86.84	81.04	63.64	65.54	52.06	27.28
	50	T^{*}_{SLR}	97.28	91.30	71.28	85.48	70.74	34.94	91.82	80.98	48.06	71.12	53.58	19.58
	50	T^{r*}_{SLR}	94.36	86.34	63.20	80.26	64.88	28.98	81.08	66.96	36.78	61.32	44.54	16.94
	100	T_{SLR}	100.0	99.98	99.98	99.30	97.66	88.90	100.0	100.0	99.96	96.96	92.62	73.16
	100	T^{r}_{SLR}	97.60	97.08	95.56	96.14	94.50	85.62	98.08	97.78	96.42	96.90	93.40	72.38
	100	T^{*}_{SLR}	99.34	96.76	85.28	93.34	85.22	61.02	99.52	97.00	84.20	96.02	87.58	53.88
	100	T^{r*}_{SLR}	98.38	95.14	81.50	89.32	80.36	55.42	98.28	94.26	77.56	90.92	80.70	52.68
	200	T_{SLR}	100.0	100.0	100.0	100.0	99.96	99.50	100.0	100.0	100.0	99.98	99.94	99.28
	200	T^{r}_{SLR}	99.36	99.20	99.06	99.56	99.42	98.24	99.32	99.20	98.76	99.50	99.24	97.64
	200	T^{*}_{SLR}	100.0	100.0	99.14	98.40	94.78	79.06	100.0	99.94	98.46	99.02	96.66	85.70
	200	T^{r*}_{SLR}	99.98	99.90	97.56	96.38	91.40	72.80	97.94	96.86	91.94	89.88	84.94	70.42
	500	T_{SLR}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	500	T^{r}_{SLR}	100.0	100.0	100.0	100.0	100.0	99.98	100.0	100.0	100.0	100.0	100.0	100.0
	500	T^{*}_{SLR}	100.0	99.94	99.54	99.48	98.40	93.18	100.0	100.0	100.0	99.96	99.48	97.34
	500	T^{r*}_{SLR}	99.96	99.92	99.12	98.92	97.08	89.78	95.70	95.40	94.58	92.32	90.94	86.00
(0.2,0.8)	50	T_{SLR}	97.72	95.04	83.52	76.86	64.76	39.46	95.80	92.04	77.54	71.14	59.34	34.02
	50	T^{r}_{SLR}	91.44	87.76	72.90	73.54	62.68	38.42	92.14	87.20	65.50	67.24	54.78	31.14
	50	T^{*}_{SLR}	88.54	73.06	36.08	70.52	53.08	20.94	88.38	74.84	37.92	69.96	51.82	21.50
	50	T^{r*}_{SLR}	61.28	64.02	30.18	64.12	46.64	18.14	79.28	62.72	28.74	59.22	43.14	16.18
	100	T_{SLR}	100.0	100.0	99.98	98.18	95.66	80.88	100.0	100.0	99.96	97.16	92.80	73.06
	100	T^{r}_{SLR}	97.52	96.98	95.28	95.98	93.24	79.96	99.66	99.62	98.90	97.62	94.88	74.60
	100	T^{*}_{SLR}	99.18	95.82	76.38	93.30	85.50	57.40	99.48	97.54	84.72	97.54	91.06	58.36
	100	T^{r*}_{SLR}	96.80	91.62	70.22	87.66	78.94	51.28	97.90	93.48	75.42	90.96	81.26	49.68
	200	T_{SLR}	100.0	100.0	100.0	99.92	99.80	98.72	100.0	100.0	100.0	99.94	99.92	98.86
	200	T^{r}_{SLR}	99.60	99.42	99.22	99.30	99.08	97.40	99.96	99.96	99.88	99.84	99.68	98.24
	200	T^{*}_{SLR}	100.0	99.96	98.86	99.38	97.46	89.18	99.92	99.32	95.28	98.80	96.10	85.66
	200	T^{r*}_{SLR}	99.80	99.56	97.42	97.30	94.06	82.44	98.54	95.54	82.10	93.46	85.92	65.04
	500	T_{SLR}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	500	T^{r}_{SLR}	99.98	99.98	99.98	99.98	99.98	99.84	100.0	100.0	100.0	100.0	100.0	99.98
	500	T^{*}_{SLR}	99.80	98.84	94.82	98.58	96.48	91.68	100.0	99.98	99.46	99.64	98.64	94.08
	500	T^{r*}_{SLR}	99.60	98.58	94.42	98.34	95.88	89.42	99.76	99.50	97.04	97.14	94.16	83.08
(0.8,0.2)	50	T_{SLR}	99.60	98.76	93.20	87.20	76.98	37.72	99.38	98.46	94.76	88.06	79.18	51.46
	50	T^{r}_{SLR}	93.20	91.00	83.00	82.70	72.50	45.54	93.16	90.70	83.50	82.26	73.14	51.10
	50	T^{*}_{SLR}	97.66	93.06	70.54	88.18	74.92	35.94	98.12	93.92	74.30	89.30	79..70	44.84
	50	T^{r*}_{SLR}	93.30	86.58	59.92	81.98	67.90	32.76	93.12	85.72	62.10	81.34	69.02	33.76
	100	T_{SLR}	100.0	100.0	99.98	98.32	95 .90	84.50	100.0	100.0	99.86	95.04	89.78	66.20
	100	T^{r}_{SLR}	97.60	96.76	94.96	95.72	92.30	81.12	99.28	98.92	97.24	90.84	84.04	62.80
	100	T^{*}_{SLR}	93.42	84.12	60.08	84.74	67.74	21.04	89.56	77.68	54.28	90.48	77.76	37.94
	100	T^{r*}_{SLR}	87.70	77.90	52.92	79.06	62.60	22.60	79.54	67.28	46.40	80.00	65.36	29.00
	200	T_{SLR}	98.30	100.0	100.0	99.96	99.92	99.30	100.0	100.0	100.0	99.92	99.56	95.40
	200	T^{r}_{SLR}	98.30	98.00	97.12	98.16	97.82	96.08	98.60	98.14	96.64	98.46	97.42	90.70
	200	T^{*}_{SLR}	97.90	93.54	81.76	94.04	88.78	73.74	98.86	95.28	77.50	97.46	92.86	63.28
	200	T^{r*}_{SLR}	97.50	93.04	80.76	93.00	86.60	69.24	94.00	85.88	66.00	89.40	81.06	53.38
	500	T_{SLR}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	500	T^{r}_{SLR}	99.02	98.84	98.52	99.06	98.94	98.48	99.86	99.84	99.80	99.86	99.84	99.74
	500	T^{*}_{SLR}	99.96	99.74	98.04	98.94	97.58	93.22	100.0	100.0	100.0	99.96	99.86	98.40
	500	T^{r*}_{SLR}	99.92	99.54	97.44	98.50	96.94	91.48	99.94	99.94	99.72	99.38	98.96	96.08
(0.8,0.8)	50	T_{SLR}	98.94	97.54	89.54	87.72	76.18	41.20	98.76	97.20	88.66	77.34	63.30	36.32
	50	T^{r}_{SLR}	96.92	94.74	84.38	87.98	78.70	51.44	96.38	93.30	81.24	78.86	67.02	42.06
	50	T^{*}_{SLR}	86.38	70.96	33.28	86.30	72.96	36.32	91.56	78.70	46.46	85.60	71.28	33.58
	50	T^{r*}_{SLR}	77.48	61.74	27.54	78.06	63.72	30.24	81.16	64.66	34.98	66.90	53.42	25.22
	100	T_{SLR}	99.96	99.94	99.76	96.84	93.78	75.74	100.0	100.0	99.96	98.44	96.16	84.38
	100	T^{r}_{SLR}	99.48	99.28	98.20	96.52	92.64	75.38	99.98	99.98	99.76	98.28	96.30	86.98
	100	T^{*}_{SLR}	96.36	87.74	63.10	91.90	81.22	52.52	98.96	98.96	67.06	93.88	85.92	58.42
	100	T^{r*}_{SLR}	88.42	76.24	51.46	83.56	72.94	44.16	86.02	86.02	46.10	82.60	71.66	44.24
	200	T_{SLR}	100.0	100.0	100.0	99.96	98.78	98.00	100.0	100.0	100.0	100.0	99.92	98.84
	200	T^{r}_{SLR}	99.74	99.70	99.54	99.86	99.78	96.94	99.76	99.76	99.62	99.46	99.32	96.90
	200	T^{*}_{SLR}	99.26	96.86	85.40	96.78	91.64	75.08	98.04	93.66	81.24	97.38	93.02	77.48
	200	T^{r*}_{SLR}	96.90	93.34	80.44	92.54	86.54	68.26	92.80	86.34	70.02	90.02	80.34	55.70
	500	T_{SLR}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	500	T^{r}_{SLR}	100.0	100.0	100.0	100.0	100.0	99.84	100.0	100.0	100.0	100.0	100.0	99.96
	500	T^{*}_{SLR}	99.92	99.22	95.32	98.46	96.46	91.78	100.0	100.0	99.48	99.54	98.72	94.54
	500	T^{r*}_{SLR}	99.74	98.56	94.36	97.48	95.44	89.94	99.52	99.02	93.72	96.34	92.20	77.92

Table 3 – Empirical Size of the Tests for the FE-SPD Model, **DGP1**

			Queen contiguity						Group interaction					
			Normal Disturbances			Non-normal Disturbances			Normal Disturbances			Non-normal Disturbances		
(λ_1,λ_2)	n	Tests	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
(0.2,0.2)	50	T_{SPD}	13.08	7.02	1.74	24.56	16.82	7.20	13.58	7.68	2.02	24.22	15.80	6.50
	50	T_{SPD}^{r}	7.48	3.84	0.76	15.58	9.38	3.30	10.24	5.16	1.08	18.28	10.48	3.00
	50	T_{SPD}^{rr}	7.20	3.54	0.60	9.94	4.94	0.94	9.80	4.60	0.64	11.02	5.32	0.82
	50	T_{SPD}^{*}	10.68	5.18	0.88	18.72	10.52	2.44	11.14	5.50	1.04	18.20	9.92	1.66
	50	T_{SPD}^{r*}	10.12	4.76	0.74	17.10	9.40	2.14	10.56	5.06	0.80	16.94	8.86	1.28
	50	T_{SPD}^{rr*}	9.82	4.64	0.52	9.48	4.08	0.54	10.50	4.80	0.66	9.32	3.56	0.38
	100	T_{SPD}	11.12	5.90	1.22	21.62	14.06	5.40	11.46	6.28	1.32	22.38	14.60	5.36
	100	T_{SPD}^{r}	8.92	4.70	0.70	17.20	10.68	3.70	9.76	5.20	0.94	19.12	11.10	3.56
	100	T_{SPD}^{rr}	9.24	4.50	0.80	10.74	5.32	1.00	9.98	5.10	0.84	11.54	5.76	1.28
	100	T_{SPD}^{*}	9.50	4.76	0.80	17.30	10.34	2.82	10.06	5.18	0.84	18.78	10.30	2.64
	100	T_{SPD}^{r*}	9.32	4.64	0.80	16.56	9.74	2.60	9.84	4.82	0.78	18.02	9.60	2.44
	100	T_{SPD}^{rr*}	9.64	4.54	0.84	9.42	4.22	0.60	10.08	4.82	0.78	9.70	4.28	0.68
	200	T_{SPD}	10.68	5.66	1.02	21.22	13.42	4.98	10.04	5.00	1.10	20.52	13.44	5.26
	200	T_{SPD}^{r}	10.04	5.04	0.80	19.42	11.48	3.98	9.46	4.48	0.92	18.66	11.78	4.16
	200	T_{SPD}^{rr}	10.08	5.06	0.78	11.40	6.16	1.04	9.38	4.32	0.86	11.68	6.44	1.18
	200	T_{SPD}^{*}	10.00	4.88	0.82	18.78	10.28	3.32	9.36	4.38	0.74	17.94	11.06	3.22
	200	T_{SPD}^{r*}	9.88	4.80	0.80	18.34	10.28	3.14	9.12	4.34	0.74	17.66	10.74	3.12
	200	T_{SPD}^{rr*}	9.98	4.94	0.70	10.24	5.38	0.72	9.06	4.16	0.64	10.36	5.34	0.74
	500	T_{SPD}	10.38	5.18	1.16	19.00	11.82	3.94	11.12	5.62	1.16	19.10	12.26	4.10
	500	T_{SPD}^{r}	9.96	4.90	1.08	18.04	11.08	3.28	10.42	5.00	1.06	17.82	10.94	3.38
	500	T_{SPD}^{rr}	9.90	5.12	1.02	10.54	5.44	0.94	10.28	5.00	1.00	10.74	4.98	1.10
	500	T_{SPD}^{*}	9.94	4.82	1.08	17.78	10.44	2.84	10.80	5.34	1.10	17.94	11.14	3.22
	500	T_{SPD}^{r*}	9.84	4.74	1.08	17.54	10.22	2.78	10.76	5.30	1.06	17.86	11.04	3.12
	500	T_{SPD}^{rr*}	9.86	5.00	1.00	9.64	4.90	0.74	10.68	5.24	0.88	10.54	4.80	0.88
(0.2,0.8)	50	T_{SPD}	13.80	7.42	1.92	24.30	16.52	7.22	13.00	7.44	2.00	25.36	17.00	7.14
	50	T_{SPD}^{r}	10.88	5.36	1.24	19.20	12.32	4.18	11.08	5.96	1.40	20.40	12.90	4.82
	50	T_{SPD}^{rr}	10.70	5.26	0.94	12.40	6.12	1.04	11.28	5.82	1.22	13.56	7.10	1.60
	50	T_{SPD}^{*}	11.22	5.12	0.98	18.00	10.84	2.42	9.78	5.08	1.22	18.40	10.52	3.42
	50	T_{SPD}^{r*}	10.52	4.60	0.76	16.92	9.58	1.98	8.84	4.50	0.90	16.12	8.68	2.44
	50	T_{SPD}^{rr*}	10.74	5.00	0.70	9.76	3.70	0.54	8.36	3.84	0.50	9.20	4.46	0.64
	100	T_{SPD}	12.02	6.94	1.82	22.28	14.44	5.12	11.76	5.84	1.08	22.34	14.92	5.78
	100	T_{SPD}^{r}	10.68	5.82	1.44	19.24	11.76	3.36	10.38	5.16	0.82	19.64	12.06	3.88
	100	T_{SPD}^{rr}	10.90	5.54	1.16	11.58	5.46	1.14	10.32	4.78	0.90	11.94	5.94	1.14
	100	T_{SPD}^{*}	10.84	5.66	1.34	18.84	10.78	2.54	10.28	4.76	0.64	19.06	10.92	2.84
	100	T_{SPD}^{r*}	10.60	5.54	1.28	17.98	9.84	2.28	9.44	4.26	0.42	17.44	9.24	2.28
	100	T_{SPD}^{rr*}	10.92	5.28	1.04	9.82	4.20	0.64	9.20	3.98	0.52	9.48	4.02	0.68
	200	T_{SPD}	10.24	5.04	1.00	20.46	12.56	4.46	11.10	5.82	1.18	20.42	12.60	4.36
	200	T_{SPD}^{r}	9.56	4.60	0.88	18.38	10.96	3.56	10.68	5.36	1.08	18.68	11.00	3.28
	200	T_{SPD}^{rr}	9.54	4.62	0.80	10.68	5.76	1.12	10.56	5.16	1.04	11.12	5.28	1.02
	200	T_{SPD}^{*}	9.62	4.48	0.80	18.28	10.16	2.92	10.10	5.22	0.92	17.78	10.12	2.72
	200	T_{SPD}^{r*}	9.32	4.36	0.80	17.68	9.92	2.74	9.58	4.72	0.80	16.74	9.10	2.12
	200	T_{SPD}^{rr*}	9.46	4.50	0.74	9.54	4.80	0.56	9.40	4.30	0.74	9.41	4.50	0.64
	500	T_{SPD}	10.86	5.98	1.06	18.64	11.58	3.92	10.42	5.42	1.22	19.14	12.00	4.44
	500	T_{SPD}^{r}	10.56	5.80	0.98	17.82	10.88	3.58	10.18	5.26	1.16	18.30	11.00	3.80
	500	T_{SPD}^{rr}	10.22	5.66	1.04	10.40	5.48	1.00	10.12	5.36	1.18	10.56	5.62	1.32
	500	T_{SPD}^{*}	10.40	5.68	0.90	17.38	10.44	3.28	10.20	5.14	1.16	18.06	10.58	3.42
	500	T_{SPD}^{r*}	10.36	5.66	0.90	17.14	10.30	3.02	10.00	5.04	1.10	17.66	10.28	3.16
	500	T_{SPD}^{rr*}	10.06	5.50	0.96	9.90	4.78	0.86	9.96	5.04	0.98	9.92	4.94	0.98
(0.8,0.8)	50	T_{SPD}	12.68	6.70	1.80	25.32	16.98	6.52	14.44	7.70	2.00	23.38	16.06	6.22
	50	T_{SPD}^{r}	10.80	5.48	1.28	20.98	12.80	3.48	12.64	6.14	1.48	19.74	12.22	3.86
	50	T_{SPD}^{rr}	10.64	5.32	0.72	12.68	6.04	0.82	12.62	5.76	0.96	12.64	5.96	1.12
	50	T_{SPD}^{*}	1.00	4.96	0.84	18.48	9.84	2.00	11.92	5.62	1.00	17.86	10.12	2.24
	50	T_{SPD}^{r*}	9.32	4.54	0.64	17.22	8.86	1.60	10.76	5.00	0.86	16.10	8.44	1.62
	50	T_{SPD}^{rr*}	9.88	4.52	0.46	9.28	3.56	0.51	10.70	4.64	0.62	8.88	3.18	0.40
	100	T_{SPD}	11.24	5.96	1.16	22.42	14.90	6.02	11.94	6.56	1.62	23.36	15.60	6.24
	100	T_{SPD}^{r}	10.24	5.16	0.90	19.70	12.32	4.18	11.10	5.84	1.38	20.52	13.08	4.12
	100	T_{SPD}^{rr}	10.60	5.16	0.84	12.50	6.70	1.32	11.00	5.72	1.14	12.56	6.68	1.50
	100	T_{SPD}^{*}	10.02	4.82	0.74	18.74	10.94	2.90	10.54	5.26	1.04	19.40	11.54	2.86
	100	T_{SPD}^{r*}	9.80	4.58	0.68	17.88	10.44	2.54	9.68	4.60	0.86	17.20	9.86	2.22
	100	T_{SPD}^{rr*}	10.04	4.64	0.70	10.44	4.52	0.58	9.80	4.44	0.68	9.70	3.98	0.46
	200	T_{SPD}	10.34	5.18	0.92	19.52	12.54	4.54	10.72	5.70	1.36	20.58	13.68	5.06
	200	T_{SPD}^{r}	9.62	4.74	0.74	17.88	10.62	3.36	10.16	5.24	1.24	18.92	12.40	3.94
	200	T_{SPD}^{rr}	9.92	4.82	0.76	10.50	5.54	1.14	9.96	5.44	1.24	11.92	6.36	1.48
	200	T_{SPD}^{*}	9.28	4.54	0.66	17.12	9.62	2.70	10.00	4.98	0.98	18.30	11.46	3.26
	200	T_{SPD}^{r*}	9.08	4.32	0.60	16.64	9.16	2.54	9.54	4.70	0.90	17.04	10.24	2.80
	200	T_{SPD}^{rr*}	9.64	4.48	0.66	9.18	4.52	0.50	9.32	4.50	0.86	9.60	4.50	0.80
	500	T_{SPD}	11.30	5.88	1.20	18.64	11.72	3.60	10.92	5.66	1.22	18.90	11.86	4.00
	500	T_{SPD}^{r}	11.02	5.68	1.12	17.76	10.96	3.22	10.70	5.54	1.10	17.94	11.20	3.44
	500	T_{SPD}^{rr}	11.30	5.76	1.08	10.32	4.92	0.86	10.80	5.38	1.22	10.54	5.08	1.00
	500	T_{SPD}^{*}	11.00	5.58	1.06	17.54	10.60	2.82	10.58	5.32	1.10	17.74	10.86	2.98
	500	T_{SPD}^{r*}	10.90	5.56	1.06	17.32	10.42	2.80	10.30	5.14	1.06	17.30	10.28	2.70
	500	T_{SPD}^{rr*}	11.12	5.68	1.04	9.82	4.30	0.70	10.32	5.14	1.06	9.80	4.38	0.72

Table 4 – Empirical Size Adjusted Powers of the Tests for the FE-SPD Model, **DGP1 – Disturbances Heteroskedastic $\alpha = 0.5$**

			Queen contiguity						Group interaction					
			Normal Disturbances			Non-normal Disturbances			Normal Disturbances			Non-normal Disturbances		
(λ_1,λ_2)	n	Tests	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
(0.2,0.2)	50	T_{SPD}	90.14	82.42	60.04	50.74	38.14	16.42	74.10	62.60	37.60	52.70	38.68	16.94
	50	T_{SPD}^{r}	79.34	70.60	48.00	46.34	33.92	15.38	72.70	60.66	35.58	52.28	38.94	16.66
	50	T_{SPD}^{rr}	77.70	66.62	40.88	44.42	32.38	14.48	71.40	59.30	31.74	52.18	37.60	15.88
	50	T_{SPD}^{*}	88.20	78.90	51.36	55.78	43.40	20.74	64.96	48.04	15.46	50.28	34.86	13.42
	50	T_{SPD}^{r*}	87.80	78.86	49.68	55.82	42.66	19.88	63.92	47.06	14.18	48.96	33.50	12.58
	50	T_{SPD}^{rr*}	84.56	71.00	41.70	52.70	39.78	18.14	62.30	43.26	12.78	47.84	31.92	11.18
	100	T_{SPD}	95.66	91.56	76.46	71.34	58.78	30.42	96.48	92.82	79.56	69.06	55.76	31.14
	100	T_{SPD}^{r}	95.18	90.66	74.02	71.32	58.76	31.10	95.94	91.96	76.26	68.90	57.10	29.38
	100	T_{SPD}^{rr}	94.46	89.50	69.34	70.44	56.40	28.84	95.20	90.74	70.18	68.80	55.78	37.64
	100	T_{SPD}^{*}	95.14	89.64	68.20	69.46	54.82	24.36	95.36	89.10	67.76	67.02	52.96	22.82
	100	T_{SPD}^{r*}	95.06	89.34	68.00	69.12	54.62	24.54	95.30	88.86	67.26	67.00	51.90	22.70
	100	T_{SPD}^{rr*}	93.98	86.82	60.28	67.98	51.52	23.36	93.84	86.24	58.72	66.22	50.00	21.62
	200	T_{SPD}	99.94	99.92	99.24	95.16	91.36	75.60	100.0	99.96	99.50	93.56	87.72	68.46
	200	T_{SPD}^{r}	99.68	99.56	98.68	94.82	90.74	75.02	99.90	99.86	99.30	93.54	87.66	68.94
	200	T_{SPD}^{rr}	99.70	99.52	98.18	94.54	89.98	72.84	99.90	99.86	99.12	93.02	86.46	65.82
	200	T_{SPD}^{*}	99.92	99.84	98.80	95.76	91.54	75.06	100.0	99.96	99.14	93.82	87.88	68.08
	200	T_{SPD}^{r*}	99.92	99.84	98.82	95.62	91.42	74.86	100.0	99.96	99.12	93.70	87.62	67.70
	200	T_{SPD}^{rr*}	99.92	99.76	97.80	95.26	90.08	72.18	100.0	99.94	98.44	93.06	86.54	63.30
	500	T_{SPD}	100.0	100.0	100.0	99.96	99.90	98.98	100.0	100.0	100.0	100.0	99.88	99.38
	500	T_{SPD}^{r}	100.0	100.0	100.0	99.96	99.90	99.02	99.98	99.98	99.98	99.96	99.84	99.30
	500	T_{SPD}^{rr}	100.0	100.0	100.0	99.96	99.80	98.68	99.98	99.98	99.98	99.96	99.84	99.16
	500	T_{SPD}^{*}	100.0	100.0	100.0	99.96	99.90	98.70	100.0	100.0	100.0	99.98	99.90	99.32
	500	T_{SPD}^{r*}	100.0	100.0	100.0	99.96	99.90	98.68	100.0	100.0	100.0	99.98	99.90	99.22
	500	T_{SPD}^{rr*}	100.0	100.0	100.0	99.96	99.76	98.08	100.0	100.0	100.0	100.0	99.84	99.04
(0.2,0.8)	50	T_{SPD}	63.68	49.44	23.04	39.16	27.30	11.00	60.98	48.10	24.06	37.56	26.62	17.94
	50	T_{SPD}^{r}	61.56	45.64	19.28	38.48	26.20	10.51	59.06	45.30	22.32	37.66	27.44	16.66
	50	T_{SPD}^{rr}	60.40	46.00	19.20	38.24	26.48	10.43	57.52	43.36	20.16	35.60	24.58	14.16
	50	T_{SPD}^{*}	61.06	43.00	15.14	44.16	32.82	13.60	55.12	38.68	14.28	37.24	25.96	12.40
	50	T_{SPD}^{r*}	60.40	42.36	14.02	43.76	31.60	12.80	54.28	37.58	12.32	36.92	24.58	11.36
	50	T_{SPD}^{rr*}	57.08	40.08	12.56	42.22	31.18	12.76	49.90	33.36	12.02	34.90	22.36	10.44
	100	T_{SPD}	96.74	93.44	79.78	71.54	59.14	30.80	97.08	93.62	80.66	73.66	61.72	36.08
	100	T_{SPD}^{r}	95.74	92.52	78.14	71.36	58.48	30.68	97.06	93.48	79.62	73.88	61.78	35.64
	100	T_{SPD}^{rr}	95.50	91.72	74.46	71.62	57.46	30.56	96.44	91.74	76.58	72.54	59.04	33.88
	100	T_{SPD}^{*}	96.18	91.36	70.16	72.62	59.32	30.70	96.80	93.48	77.72	75.88	64.62	38.72
	100	T_{SPD}^{r*}	95.98	91.46	69.86	72.06	58.74	30.50	96.76	92.74	75.88	74.52	62.48	34.00
	100	T_{SPD}^{rr*}	95.42	89.78	61.66	71.50	58.02	30.76	95.90	90.00	70.70	72.58	59.86	29.86
	200	T_{SPD}	99.92	99.80	99.00	93.58	87.94	72.46	99.90	99.66	97.88	90.28	83.10	60.60
	200	T_{SPD}^{r}	99.88	99.74	98.88	94.08	88.40	73.26	99.90	99.64	97.58	90.72	82.70	58.96
	200	T_{SPD}^{rr}	99.84	99.74	97.98	93.54	88.00	70.84	99.88	99.48	96.36	90.06	80.52	55.48
	200	T_{SPD}^{*}	99.92	99.76	98.18	94.04	88.06	70.86	99.90	99.40	95.48	89.86	79.68	48.16
	200	T_{SPD}^{r*}	99.90	99.76	98.16	94.06	87.96	70.90	99.44	98.10	90.16	86.52	73.90	39.26
	200	T_{SPD}^{rr*}	99.84	99.68	96.80	93.46	87.68	68.58	98.46	95.82	80.80	84.28	70.46	36.48
	500	T_{SPD}	100.0	100.0	100.0	99.94	99.78	98.62	100.0	100.0	100.0	99.98	99.88	98.62
	500	T_{SPD}^{r}	100.0	100.0	100.0	99.94	99.80	98.64	100.0	100.0	100.0	99.98	99.88	98.84
	500	T_{SPD}^{rr}	100.0	100.0	100.0	99.92	99.74	98.24	100.0	100.0	100.0	99.98	99.80	98.54
	500	T_{SPD}^{*}	100.0	100.0	100.0	99.94	99.80	98.48	100.0	100.0	100.0	99.98	99.88	98.86
	500	T_{SPD}^{r*}	100.0	100.0	100.0	99.94	99.78	98.50	99.90	99.88	99.80	99.48	99.04	96.88
	500	T_{SPD}^{rr*}	100.0	100.0	100.0	99.92	99.72	98.10	99.88	99.86	99.64	99.36	98.88	96.12
(0.8,0.8)	50	T_{SPD}	71.90	59.24	32.62	53.20	40.70	18.80	72.66	61.54	37.10	51.62	39.50	17.58
	50	T_{SPD}^{r}	70.22	57.22	29.84	53.94	40.70	18.16	72.42	60.96	35.30	53.42	41.64	19.06
	50	T_{SPD}^{rr}	69.06	55.50	25.84	53.20	40.46	18.74	70.50	57.70	30.94	51.58	38.30	16.88
	50	T_{SPD}^{*}	61.54	44.24	22.22	52.06	36.32	14.38	69.76	55.72	28.26	50.16	36.68	15.26
	50	T_{SPD}^{r*}	61.14	42.88	21.30	52.10	35.06	13.48	69.54	55.04	29.88	48.94	35.62	14.20
	50	T_{SPD}^{rr*}	58.04	37.78	19.40	49.56	35.30	13.08	66.38	51.94	23.40	47.24	33.42	12.48
	100	T_{SPD}	98.84	97.60	88.12	82.00	71.14	45.58	98.10	95.98	86.44	72.80	60.66	34.32
	100	T_{SPD}^{r}	98.76	97.24	85.92	82.30	70.88	45.38	98.00	95.92	85.52	73.16	60.58	34.62
	100	T_{SPD}^{rr}	98.28	95.46	77.84	80.72	69.18	42.68	97.84	94.92	80.56	72.26	59.38	30.20
	100	T_{SPD}^{*}	98.40	95.50	73.46	80.16	66.50	34.52	97.44	94.48	78.06	74.08	61.54	36.04
	100	T_{SPD}^{r*}	98.28	95.40	72.40	79.66	65.06	33.16	97.16	93.76	76.12	72.22	58.28	31.44
	100	T_{SPD}^{rr*}	97.28	91.50	58.54	77.64	62.42	29.76	96.36	90.90	64.90	69.86	57.02	27.30
	200	T_{SPD}	99.88	99.76	98.64	92.00	86.22	66.32	99.88	99.64	98.32	93.62	88.48	68.98
	200	T_{SPD}^{r}	99.92	99.74	98.58	92.20	86.38	65.28	99.86	99.62	98.24	93.62	88.40	68.56
	200	T_{SPD}^{rr}	99.88	99.74	98.10	91.84	85.54	64.28	99.86	99.56	97.84	93.26	87.28	66.38
	200	T_{SPD}^{*}	99.90	99.70	97.62	92.02	85.04	61.70	99.76	99.46	97.34	93.48	86.66	62.10
	200	T_{SPD}^{r*}	99.90	99.68	97.58	91.94	84.68	61.62	99.28	98.70	95.52	92.16	83.80	56.44
	200	T_{SPD}^{rr*}	99.82	99.67	96.62	91.64	84.18	60.08	99.14	98.20	92.86	91.10	82.10	52.02
	500	T_{SPD}	100.0	100.0	100.0	99.90	99.64	98.22	100.0	100.0	100.0	99.82	99.52	97.96
	500	T_{SPD}^{r}	100.0	100.0	100.0	99.88	99.60	98.34	100.0	100.0	100.0	99.82	99.52	97.96
	500	T_{SPD}^{rr}	100.0	100.0	100.0	99.82	99.58	98.02	100.0	100.0	100.0	99.78	99.32	97.38
	500	T_{SPD}^{*}	100.0	100.0	100.0	99.88	99.62	97.94	100.0	100.0	100.0	99.86	99.56	98.12
	500	T_{SPD}^{r*}	100.0	100.0	100.0	99.88	99.62	97.94	100.0	100.0	100.0	98.32	97.64	94.76
	500	T_{SPD}^{rr*}	100.0	100.0	100.0	99.84	99.64	97.56	100.0	100.0	100.0	98.22	97.44	93.72

Table 5 – Empirical Size and Size adjusted Powers of the Tests for the FE-DSPD Model, $\rho = 0.3$, **DGP1**

			Queen contiguity						Group interaction					
			Normal Disturbances			Non-normal Disturbances			Normal Disturbances			Non-normal Disturbances		
$(\lambda_1, \lambda_2, \lambda_3)$	n	Tests	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
Empirical Size														
(0.2,0.2,0.2)	50	T^{r}_{DSPD}	10.48	4.78	0.76	10.68	5.36	0.54	11.68	5.68	0.86	12.14	6.04	1.20
	50	T^{r*}_{DSPD}	9.40	3.90	0.50	9.58	4.54	0.48	10.50	5.06	0.42	8.72	3.40	0.28
	100	T^{r}_{DSPD}	10.56	5.42	0.92	11.26	5.92	1.12	11.04	5.38	0.92	12.18	6.22	1.12
	100	T^{r*}_{DSPD}	9.86	4.78	0.74	10.74	5.40	0.90	10.36	4.72	0.76	10.04	4.52	0.44
	200	T^{r}_{DSPD}	10.24	5.20	1.02	10.22	4.62	0.68	11.02	5.72	1.08	12.04	6.50	1.06
	200	T^{r*}_{DSPD}	10.00	5.02	0.90	9.96	4.34	0.58	10.64	5.38	0.88	10.92	5.24	0.72
	400	T^{r}_{DSPD}	10.32	5.16	0.84	11.00	5.04	1.12	10.88	5.30	1.10	11.02	5.58	1.02
	400	T^{r*}_{DSPD}	10.08	5.02	0.80	10.86	4.86	0.96	10.64	5.18	1.04	10.26	4.88	0.80
(-0.2,-0.2,-0.2)	50	T^{r}_{DSPD}	11.52	5.66	0.98	13.94	7.42	1.46	12.30	6.00	1.10	13.64	6.98	1.16
	50	T^{r*}_{DSPD}	10.30	4.80	0.68	10.50	4.62	0.54	10.96	4.94	0.76	10.24	4.28	0.32
	100	T^{r}_{DSPD}	11.70	5.78	1.02	12.66	6.24	1.36	11.30	5.64	0.94	12.24	6.18	1.16
	100	T^{r*}_{DSPD}	10.94	5.34	0.72	10.36	4.36	0.56	10.66	4.86	0.64	10.22	4.46	0.42
	200	T^{r}_{DSPD}	10.76	5.32	1.10	11.48	5.74	1.12	9.76	5.26	1.00	11.52	5.36	1.20
	200	T^{r*}_{DSPD}	10.36	4.04	0.98	10.22	4.64	0.58	9.56	4.96	0.92	10.30	4.42	0.72
	400	T^{r}_{DSPD}	9.84	5.02	1.02	10.92	5.22	0.74	10.30	5.02	0.94	10.74	5.36	1.10
	400	T^{r*}_{DSPD}	9.62	4.92	0.94	10.36	4.64	0.64	10.12	4.92	0.84	10.00	4.70	0.66
Size adjusted Powers - Disturbances Heteroskedastic $\alpha = 0.5$														
(0.2,0.2,0.2)	50	T^{r}_{DSPD}	99.56	99.10	93.28	92.08	84.48	59.66	94.34	93.82	91.68	89.52	86.08	71.18
	50	T^{r*}_{DSPD}	99.38	97.54	81.24	89.68	77.96	45.92	97.42	96.90	91.60	92.58	86.18	66.90
	100	T^{r}_{DSPD}	100.0	100.0	100.0	99.92	99.46	96.16	99.82	99.82	99.82	99.42	99.04	95.88
	100	T^{r*}_{DSPD}	100.0	100.0	99.96	99.78	98.64	89.62	99.88	99.86	99.36	99.60	98.04	89.10
	200	T^{r}_{DSPD}	100.0	100.0	100.0	100.0	99.98	99.12	99.64	99.64	99.62	99.26	99.24	98.72
	200	T^{r*}_{DSPD}	100.0	100.0	99.92	99.84	99.38	96.82	99.92	99.92	99.90	99.92	99.70	98.14
	400	T^{r}_{DSPD}	100.0	100.0	100.0	100.0	100.0	99.92	100.0	100.0	100.0	100.0	100.0	99.98
	400	T^{r*}_{DSPD}	100.0	100.0	100.0	99.98	99.90	99.06	100.0	100.0	100.0	100.0	99.96	99.78
(-0.2,-0.2,-0.2)	50	T^{r}_{DSPD}	99.76	99.46	95.52	93.96	88.26	63.28	99.46	99.38	97.82	95.28	91.74	76.42
	50	T^{r*}_{DSPD}	99.54	97.76	86.66	91.48	81.42	49.40	99.92	99.60	94.96	94.80	88.70	70.40
	100	T^{r}_{DSPD}	100.0	100.0	100.0	99.70	99.38	96.74	100.0	100.0	99.98	99.78	99.26	95.48
	100	T^{r*}_{DSPD}	100.0	100.0	99.96	99.62	98.68	91.92	100.0	100.0	99.66	99.44	97.68	89.12
	200	T^{r}_{DSPD}	100.0	100.0	100.0	100.0	100.0	99.34	99.98	99.98	99.98	99.98	99.96	99.50
	200	T^{r*}_{DSPD}	100.0	100.0	99.84	99.88	99.48	97.58	100.0	100.0	100.0	99.92	99.80	98.38
	400	T^{r}_{DSPD}	100.0	100.0	100.0	100.0	100.0	99.96	100.0	100.0	100.0	100.0	100.0	99.98
	400	T^{r*}_{DSPD}	100.0	100.0	100.0	100.0	99.98	99.62	100.0	100.0	100.0	100.0	99.98	99.90

Table 6 – Empirical Size and Size adjusted Powers of the Tests for the FE-DSPD Model, $\rho = 0.3$, **DGP2**

			Queen contiguity						Group interaction					
			Normal Disturbances			Non-normal Disturbances			Normal Disturbances			Non-normal Disturbances		
$(\lambda_1, \lambda_2, \lambda_3)$	n	Tests	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
Empirical Size														
(0.2,0.2,0.2)	50	T^{r}_{DSPD}	13.74	6.94	1.34	13.56	7.12	1.18	11.92	6.36	1.14	11.92	6.36	1.14
	50	T^{r*}_{DSPD}	10.12	4.44	0.20	10.26	4.34	0.36	8.78	3.62	0.38	8.78	3.62	0.38
	100	T^{r}_{DSPD}	12.70	6.62	1.08	12.34	6.20	1.38	11.72	6.02	1.26	11.72	6.02	1.26
	100	T^{r*}_{DSPD}	10.48	4.54	0.56	10.12	4.60	0.48	9.60	4.36	0.50	9.60	4.36	0.50
	200	T^{r}_{DSPD}	12.00	6.38	1.06	12.24	6.12	1.14	11.54	5.56	1.12	11.54	5.56	1.12
	200	T^{r*}_{DSPD}	10.72	5.06	0.60	10.68	5.18	0.72	10.12	4.54	0.60	10.12	4.54	0.60
	400	T^{r}_{DSPD}	11.18	5.34	1.02	10.74	5.42	1.12	11.00	5.80	1.20	11.00	5.80	1.20
	400	T^{r*}_{DSPD}	10.64	4.98	0.76	10.02	4.86	0.76	10.36	4.88	0.82	10.36	4.88	0.82
(-0.2,-0.2,-0.2)	50	T^{r}_{DSPD}	10.94	5.42	0.92	14.52	7.94	1.74	11.56	5.40	1.10	14.74	7.32	1.54
	50	T^{r*}_{DSPD}	9.76	4.44	0.68	10.56	4.50	0.58	10.14	4.30	0.52	10.24	4.08	0.46
	100	T^{r}_{DSPD}	11.00	5.68	1.10	12.86	6.86	1.22	10.98	5.04	1.06	11.92	6.40	1.08
	100	T^{r*}_{DSPD}	10.40	5.04	0.72	10.64	4.70	0.64	10.46	4.56	0.74	9.72	4.36	0.44
	200	T^{r}_{DSPD}	10.68	5.44	0.94	11.64	5.96	1.18	10.12	4.86	0.98	11.30	5.40	1.10
	200	T^{r*}_{DSPD}	10.18	5.28	0.86	10.60	4.72	0.64	9.88	4.54	0.76	10.18	4.38	0.70
	400	T^{r}_{DSPD}	10.86	5.30	1.10	10.62	5.44	1.14	9.86	5.06	1.16	10.74	5.02	1.04
	400	T^{r*}_{DSPD}	10.76	5.18	1.02	10.02	4.98	0.76	9.72	4.84	1.06	9.84	4.64	0.82
Size adjusted Powers - Disturbances Heteroskedastic $\alpha = 0.5$														
(0.2,0.2,0.2)	50	T^{r}_{DSPD}	99.10	98.66	93.52	91.80	84.00	58.00	86.20	85.08	80.10	82.40	77.30	56.72
	50	T^{r*}_{DSPD}	99.12	97.14	85.68	89.72	80.58	53.54	95.28	90.26	70.56	89.44	80.02	46.94
	100	T^{r}_{DSPD}	99.44	99.44	99.36	98.84	98.44	92.56	94.42	94.38	94.24	93.06	92.84	91.66
	100	T^{r*}_{DSPD}	99.78	99.78	99.10	98.68	96.88	86.92	97.22	96.98	95.98	96.42	95.46	90.42
	200	T^{r}_{DSPD}	100.0	100.0	100.0	99.98	99.96	99.88	99.64	99.60	99.56	99.44	99.34	98.62
	200	T^{r*}_{DSPD}	100.0	100.0	99.92	99.96	99.82	98.86	100.0	100.0	99.94	99.92	99.42	97.38
	400	T^{r}_{DSPD}	100.0	100.0	100.0	100.0	100.0	99.98	99.96	99.96	99.96	99.92	99.88	99.88
	400	T^{r*}_{DSPD}	100.0	100.0	100.0	100.0	99.96	99.78	100.0	100.0	100.0	100.0	99.88	99.92
(-0.2,-0.2,-0.2)	50	T^{r}_{DSPD}	99.64	99.00	90.98	92.64	84.72	60.96	93.56	92.22	87.32	89.52	83.82	62.54
	50	T^{r*}_{DSPD}	99.38	97.24	80.92	90.34	80.90	56.38	97.92	92.96	74.80	91.92	83.08	49.08
	100	T^{r}_{DSPD}	99.74	99.74	99.62	99.50	98.88	92.40	99.90	99.90	99.90	99.84	99.80	98.68
	100	T^{r*}_{DSPD}	99.94	99.92	99.02	99.14	97.50	85.46	99.90	99.90	99.64	99.76	99.22	94.02
	200	T^{r}_{DSPD}	100.0	100.0	100.0	99.98	99.98	99.68	100.0	100.0	100.0	100.0	100.0	99.22
	200	T^{r*}_{DSPD}	100.0	100.0	100.0	99.94	99.56	98.22	100.0	100.0	99.90	99.90	99.30	96.46
	400	T^{r}_{DSPD}	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	99.98	99.98	99.96
	400	T^{r*}_{DSPD}	100.0	100.0	100.0	100.0	100.0	99.82	100.0	100.0	100.0	100.0	100.0	99.96