# Joint Tests for Dynamic and Spatial Effects in Short Panel Data Models with Fixed Effects

Zhenlin Yang*
*School of Economics, Singapore Management University*
Email: zlyang@smu.edu.sg

February 28, 2016

**Abstract**

Many empirical economic research involves panel data where the time dimension $T$ is small. Of immediate interest is to test whether dynamic and spatial effects are relevant, but handy methods are unavailable. The usual LM or score tests based on the conditional likelihood given the initial observations are inconsistent when $T$ is fixed. In this paper, simple tests, referred to as *adjusted quasi score* (AQS) tests, for dynamic and/or spatial effects in fixed-effects panel data models are proposed. The AQS tests are free from specifications of the initial conditions and are consistent when $T$ is fixed. Standardized AQS tests are also derived, which are shown to have much improved finite sample properties. All the proposed tests are robust against nonnormality. Certain joint or conditional tests are fully robust against cross-sectional heteroskedasticity; the others are fairly robust against mild departures from homoskedasticity. Monte Carlo results show excellent performance of the standardized AQS tests.

**Key Words:** Adjusted quasi scores; Dynamic effect; Fixed effects; Heteroskedasticity; Initial conditions free; Nonnormality; Short panels; Tests; Spatial effects.

**JEL classifications:** C12, C18, C21, C23.

## 1. Introduction

Fixed-effects panel data (FE-PD) model has been an important tool for the applied economics researchers over the past few decades. However, there have been growing concerns on whether the panel models are dynamic in nature due to the impacts from the past and current to the future 'economic' performance, and whether the models contain spatial dependence caused by the interactions among economic agents or social actors (e.g., neighbourhood effects, copy-catting, social network, and peer group effects). In other words, there have been growing concerns from the applied researchers on whether a dynamic spatial panel data model (SDPD) with fixed-effects is more appropriate than the regular FE-PD model, or the regular

*I am grateful to Singapore Management University for financial support under Grant C244/MSS14E002.

fixed effects dynamic panel data (FE-DPD) model, or the static fixed effects spatial panel data (FE-SPD) model. Thus, it is highly desirable to device simple and reliable tests for the applied researchers to determine whether to use the complicated general FE-SDPD model or some simpler model, such as an FE-SDPD model with certain spatial effects dropped, or the static FE-SPD model with all the dynamic effects dropped, or the FE-DPD model with all spatial effects dropped, or the FE-PD model with both dynamic and spatial effects dropped.

In the dynamic panel data model literature, it is widely recognized that the (quasi) maximum likelihood (ML) type of estimation is more efficient than the GMM-type estimation (see, e.g., Hsiao, et al., 2002; Hsiao, 2003; Binder et al. 2005; Bun and Carree 2005; Gouriéroux et al. 2010; Kruiniger 2013). However, the ML-type estimation faces an important issue when the panels are *short* (there is a large number of cross-sectional units $n$ but a small number of time periods $T$): the initial observations need to be modeled to give an appropriate likelihood function so that the (quasi) ML estimator can be consistent (see, e.g., Anderson and Hsiao, 1981, 1982; Bhargava and Sargan, 1983; Hsiao, et al., 2002; and Su and Yang, 2015). Ignoring the initial conditions will result in estimators of the dynamic and spatial parameters that are inconsistent when $T$ is fixed, and have a bias of order $O(T^{-1})$ when $T$ grows with $n$ proportionally. Obviously, the inconsistency and bias in the parameter estimation will translate into the inconsistency and bias in the corresponding tests in the forms of large sample size distortion or finite sample size distortion of the tests.

While modelling the initial observations does lead to consistent and efficient ML of QML estimators when the panels are short and dynamic effect exists, it has some drawbacks: when the true model is static it may lead to estimators that have large bias and standard errors when $T$ is not so small relative to $n$, and when the model does not contain time-varying regressors it may not perform well in general. This implies the tests of lack of dynamic effect based on this estimation strategy may have a poor finite sample performance. The most serious drawback may be that this approach may not work for models containing spatial lag effects. Recently, Yang (2016) provided a unified framework for estimating a general spatial dynamic panel data (SDPD) model with fixed effects, through adjusting the conditional quasi scores (given the initial differences) associated with the dynamic and spatial parameters. He showed that the estimation based on the adjusted quasi scores leads to consistent estimators whether $T$ is fixed or grows with $n$, and that it is free from the specification of the initial conditions. Further, when $T$ grows with $n$, he showed that the new estimation method automatically corrects the bias of order $O(T^{-1})$ caused by ignoring the initial observations, and thus provides an alternative and handy method of bias correction for large panels.

However, testing problems, in particular, the joint tests for possible existence of dynamic and/or spatial effects in a panel data model have not been considered. In fact, the literature on statistical tests for the SDPD models is rather thin. This is in stark contrast with the literature on statistical tests for spatial regression models, or static spatial panel data models. See, among others, Anselin and Bera (1998), Anselin (2001), Kelejian and Prucha (2001), Yang (2010, 2015), Born and Breitung (2011), Baltagi and Yang (2013a,b), Robinson and Rossi (2014, 2015a), and Jin and Lee (2015) for spatial regression models; Baltagi et al. (2003), Baltagi et al. (2007), Debarsy and Ertur (2010), Baltagi and Yang (2013a,b), and Robinson and Rossi (2015b) for static panel data models.

In this paper, we propose simple and reliable tests for dynamic and spatial effects in fixed-effects panel data models with small $T$, which are shown to be free from the specifications of the initial conditions. The spatial effect may appear in the model in the form of spatial error (`SE`) dependence, spatial lag (`SL`) dependence, and/or space-time lag (`STL`) dependence. The initial constructions of the tests are based on the unified $M$-estimation method of Yang (2016): first adjusting the quasi score functions of the conditional quasi likelihood given the initial differences to achieve consistency, and then developing a martingale difference representation of the adjusted quasi score (AQS) function to give a consistent estimate of the variance-covariance matrix of the AQS functions. The resulted tests, referred to as AQS tests in this paper, are shown to have standard asymptotic null behavior. Further corrections are obtained on the mean and variance of the concentrated AQS functions for dynamic and spatial parameters, giving a set of standardized AQS (SAQS) tests having much better finite sample properties. All the proposed AQS and SAQS tests are robust against nonnormality. Certain joint or conditional tests are fully robust against cross-sectional heteroskedasticity; the others are fairly robust against mild departures from homoskedasticity. Monte Carlo results show excellent performance of the SAQS tests which dominate the AQS tests.

The rest of the paper is organized as follows. Section 2 introduces the general SDPD models, discusses the tests of interest, and describes the unified $M$-estimation to facilitate the construction of various tests. Section 3 presents the AQS tests and their asymptotic properties. Section 4 presents the standardized AQS tests and their asymptotic properties. Section 5 present Monte Carlo results. Section 6 concludes the paper.

## 2. Model, Tests and Unified $M$-Estimation

The spatial dynamic panel data (SDPD) model that our tests concern takes the form:

$$\begin{aligned} y_t &= \rho y_{t-1} + \lambda_1 W_1 y_t + \lambda_2 W_2 y_{t-1} + X_t\beta + Z\gamma + \mu + \alpha_t 1_n + u_t, \\ & u_t = \lambda_3 W_3 u_t + v_t, \quad t = 1, 2, \dots, T, \end{aligned} \tag{2.1}$$

where $y_t = (y_{1t}, y_{2t}, \dots, y_{nt})'$ and $v_t = (v_{1t}, v_{2t}, \dots, v_{nt})'$ are $n \times 1$ vectors of response values and idiosyncratic errors at time $t$, and $\{v_{it}\}$ are independent and identically distributed ($iid$) across $i$ and $t$ with mean zero and variance $\sigma_v^2$;[1] the scalar parameter $\rho$ characterizes the dynamic effect, $\lambda_1$ the spatial lag (`SL`) effect, $\lambda_2$ the space-time lag (`STL`) effect, and $\lambda_3$ the spatial error (`SE`) effect; $\{X_t\}$ are $n \times p$ matrices containing values of $p$ time-varying exogenous

[1]The iid assumption can be relaxed in certain cases to allow heteroskedasticity in $v_{it}$ across $i$.

variables, $Z$ is an $n \times q$ matrix containing the values of $q$ time-invariant exogenous variables that may include the constant term, dummy variables representing individuals' gender, race, etc.; $\beta$ and $\gamma$ are the usual regression coefficients; $W_r, r = 1, 2, 3$ are the given $n \times n$ spatial weight matrices; and $\mu$ is an $n \times 1$ vector of unobserved individual-specific effects, $\{\alpha_t\}$ are the time-specific effects, and $1_n$ is an $n \times 1$ vector of ones.

Model (2.1) is fairly general. It embeds several important submodels popular in the literature. Yang (2016) presented a unified, initial conditions free, method of estimation and inference for this general model, and showed that the method can easily be simplified to suit each special model of interest to a particular applied problem. A question arises naturally: in a practical application, do we really need such a general and complicated model, or does a simpler model suffice as it gives easier interpretations of the results? This suggests that before applying this general model, it is helpful to carry out some specification tests to identify a suitable model based on the data. To be exact, the tests of interest concern the dynamic and spatial parameters $\delta = (\rho, \lambda_1, \lambda_2, \lambda_3)'$. They may be the joint AQS tests (the null hypothesis sets two or more elements of $\delta$ to zero, and the 'remaining' as free parameters), or the marginal AQS tests (the null sets one element of $\delta$ to zero, and the remaining as free parameters), and the conditional AQS tests, setting some element(s) of $\delta$ to zero. Denoting $\lambda = (\lambda_1, \lambda_2, \lambda_3)'$, the specific tests of interest are as follows.

**Joint test** $H_0^{\texttt{PD}}$: *the regular panel data (PD) model suffices, that is,* $\delta = 0$.

When $H_0^{\texttt{PD}}$ is not rejected, then one proceeds with the regular panel data model and the decision is clear. However, when $H_0^{\texttt{PD}}$ is rejected, i.e., at least one element of $\delta$ is not zero, one does not know the exact cause of rejection and hence it would be necessary to carry out some sub-joint or marginal tests to identify the cause of such a rejection.

**Joint test** $H_0^{\texttt{DPD}}$: *the regular dynamic panel data (DPD) model suffices, i.e.,* $\lambda = 0$.

If $H_0^{\texttt{DPD}}$ is not rejected, then the cause of rejecting $H_0^{\texttt{PD}}$ is due to the fact that $\rho \neq 0$; otherwise, one needs to proceed with the following test.

**Marginal test** $H_0^{\texttt{STPD}}$: $\rho = 0$, *the time-space spatial panel data (STPD) model suffices.*

Under $H_0^{\texttt{STPD}}$, $\rho = 0$. Thus, if $H_0^{\texttt{STPD}}$ is not rejected, then the cause of rejecting $H_0^{\texttt{PD}}$ is that at least one element of $\lambda$ is not zero. In this case, one may proceed further to identify which element of $\lambda$ is not zero by carrying out conditional tests on one or two elements of $\lambda$, given $\rho = 0$. If $H_0^{\texttt{STPD}}$ is rejected after $H_0^{\texttt{DPD}}$ has been rejected, it is clear that at least one element of $\lambda$ is non-zero when $\rho$ is treated as a free parameter, and the marginal tests on $\lambda_r$ should be carried out, respectively, for $r = 1, 2, 3$:

**Marginal test** $H_0^{\texttt{SDPD1}}$: *the SDPD model without* $\lambda_1$ *suffices*, i.e., $H_0^{\texttt{SDPD1}}$: $\lambda_1 = 0$,
**Marginal test** $H_0^{\texttt{SDPD2}}$: *the SDPD model without* $\lambda_2$ *suffices*, i.e., $H_0^{\texttt{SDPD2}}$: $\lambda_2 = 0$,
**Marginal test** $H_0^{\texttt{SDPD3}}$: *the SDPD model without* $\lambda_3$ *suffices*, i.e., $H_0^{\texttt{SDPD3}}$: $\lambda_3 = 0$.

Note that the marginal test $H_0^{\texttt{SDPD3}}$ is quite interesting as the general model (2.1) reduces to an SDPD model with SL and STL effects under the null, which is the model considered by Lee and Yu (2008) under large $n$ and large $T$ set-up, allowing fixed individual and time effects. As seen from the following section, our test results are valid for both the cases where $T$ is fixed and where $T$ grows with $n$. The marginal test $H_0^{\texttt{SDPD2}}$ may be even more interesting as the null model becomes an SDPD model with both SL and SE (or SARAR: spatial autoregressive model with autoregressive errors) effects, which is very popular in practical applications. Another pair of joint tests of particular interest are,

**Joint test** $H_0^{\texttt{SDPD4}}$: *the SDPD model with only* SE *effect suffices, i.e.,* $\lambda_1 = \lambda_2 = 0$.
**Joint test** $H_0^{\texttt{SDPD5}}$: *the SDPD model with only* SL *effect suffices, i.e.,* $\lambda_2 = \lambda_3 = 0$.

When $H_0^{\texttt{SDPD4}}$ is true, the general model given in (2.1) reduces to an SDPD model with only the SE effect. This model is extensively studied by Su and Yang (2015) under large $n$ and small $T$ set-up, with either random or fixed individual effects. However, specification test from Model (2.1) to this reduced model has not been considered. When $H_0^{\texttt{SDPD5}}$ is true, the general model reduces to an SDPD model with only the SL effect. This is perhaps the most popular SDPD model among the applied researchers. However, a test for its adequacy in fitting the data is not available. The last test that we would like to highlight is:

**Joint test** $H_0^{\texttt{SDP}}$: *the static spatial panel data (SPD) model suffices, i.e.,* $\rho = \lambda_2 = 0$.

Under $H_0^{\texttt{SDP}}$, the model reduces to a static spatial panel data model with SL and SE (or SARAR) effects. The SARAR panel data model with fixed effects has been studied by Lee and Yu (2010) under the quasi maximum likelihood approach, LM tests for the spatial effects are given by Debarsy and Ertur (2010), and LM-type tests robust against unknown heteroskedasticity are given in Baltagi and Yang (2013b).

Besides the joint and marginal tests discussed above, some conditional tests might be of interest as well. By conditional tests we mean tests for certain types of effects, give some other effect(s) are removed from the model. For example, given that $H_0^{\texttt{SDPD2}}$ is not rejected, i.e., $\lambda_2$ is set to zero, one might be interested in testing further whether $\rho = 0$, i.e., whether the static SARAR model suffices; given that $H_0^{\texttt{STPD}}$ is not rejected, i.e., $\rho = 0$, one might be interested in testing further whether $\lambda_2 = 0$ and if so a static SARAR model suffices.

**The M-estimation.** The methodology we adopt in constructing tests statistics for testing various hypotheses discussed above closely relates to the unified $M$-estimation and inference methods presented in Yang (2016). Thus, it is necessary to outline this unified $M$-estimation method. As the current paper focuses on the fixed effects model with small $T$, the time specific effects can be absorbed into the time-varying regressors $X_t$, and the individual-specific effects need to be eliminated to avoid the incidental parameters problem. By taking the first-difference, Model (2.1) becomes,

$$\Delta y_t = \rho\Delta y_{t-1} + \lambda_1 W_1 \Delta y_t + \lambda_2 W_2 \Delta y_{t-1} + \Delta X_t \beta + \Delta u_t, \quad \Delta u_t = \lambda_3 W_3 \Delta u_t + \Delta v_t, \tag{2.2}$$

for $t = 2, 3, \cdots, T$. The parameters left in Model (2.2) are $\psi = \{\beta', \sigma_v^2, \rho, \lambda'\}'$. Note that $\Delta y_1$ depends on both the initial observations $y_0$ and the first period observations $y_1$. Thus, even if $y_0$ is exogenous, $y_1$ and thus $\Delta y_1$ is not. The $M$-estimation strategy goes as follows: formulate the conditional quasi likelihood function as if $\Delta y_1$ is exogenous, then make corrections on the relevant elements of the conditional quasi score vector, and then estimate the model parameters by solving the estimating equations defined by the adjusted quasi score functions.

Let $\Delta Y = \{\Delta y_2', \ldots, \Delta y_T'\}'$, $\Delta Y_{-1} = \{\Delta y_1', \ldots, \Delta y_{T-1}'\}'$, $\Delta X = \{\Delta X_2', \ldots, \Delta X_T'\}'$, and $\Delta v = \{\Delta v_2', \ldots, \Delta v_T'\}'$. Denote by $I_m$ an $m \times m$ identity matrix. Let $\mathbf{W}_r = I_{T-1} \otimes W_r, r = 1, 2, 3$. Let $B_r(\lambda_r) = I_n - \lambda_r W_r$, and $\mathbf{B}_r(\lambda_r) = I_{T-1} \otimes B_r(\lambda_r)$, for $r = 1$ and 3, where $\otimes$ denotes the Kronecker product. Model (2.2) can be written as:

$$\Delta Y = \rho \Delta Y_{-1} + \lambda_1 \mathbf{W}_1 \Delta Y + \lambda_2 \mathbf{W}_2 \Delta Y_{-1} + \Delta X \beta + \Delta u, \quad \Delta u = \lambda_3 \mathbf{W}_3 \Delta u + \Delta v. \tag{2.3}$$

We have, $\text{Var}(\Delta u) = \sigma_v^2 \Omega(\lambda_3)$, where $\Omega(\lambda_3) = C \otimes [B_3'(\lambda_3) B_3(\lambda_3)]^{-1}$, and,

$$C = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -1 & 2 & -1 \\ 0 & 0 & 0 & \cdots & 0 & -1 & 2 \end{pmatrix}_{(T-1)\times(T-1)}.$$

Letting $\mathbf{B}_2(\rho, \lambda_2) = \rho I_N + \lambda_2 \mathbf{W}_2$, the conditional quasi Gaussian loglikelihood of $\psi$ in terms of $\Delta y_2, \ldots, \Delta y_T$ as if $\Delta y_1$ is exogenous has the form, ignoring the constant term:

$$\ell(\psi) = -\tfrac{N}{2} \log(\sigma_v^2) - \tfrac{1}{2} \log |\Omega(\lambda_3)| + \log |\mathbf{B}_1(\lambda_1)| - \tfrac{1}{2\sigma_v^2} \Delta u(\theta)' \Omega(\lambda_3)^{-1} \Delta u(\theta), \tag{2.4}$$

where $\Delta u(\theta) = \mathbf{B}_1(\lambda_1) \Delta Y - \mathbf{B}_2(\rho, \lambda_2) \Delta Y_{-1} - \Delta X \beta$, and $\theta = (\beta', \rho, \lambda_1, \lambda_2)'$.

Maximizing $\ell(\psi)$ gives the conditional QML estimator (CQMLE) of $\psi$. Under mild conditions, maximizing the conditional loglikelihood $\ell(\psi)$ is equivalent to solving the estimating equation $S(\psi) = 0$, where $S(\psi) = \frac{\partial}{\partial \psi} \ell(\psi)$, the quasi score vector having the forms:

$$S(\psi) = \begin{cases} \frac{1}{\sigma_v^2} \Delta X' \Omega^{-1} \Delta u(\theta), \\ \frac{1}{2\sigma_v^4} \Delta u(\theta)' \Omega^{-1} \Delta u(\theta) - \frac{N}{2\sigma_v^2}, \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} \Delta Y_{-1}, \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} \mathbf{W}_1 \Delta Y - \text{tr}(\mathbf{B}_1^{-1} \mathbf{W}_1), \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} \mathbf{W}_2 \Delta Y_{-1}, \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' (C^{-1} \otimes A_3) \Delta u(\theta) - (T-1)\text{tr}(G_3), \end{cases} \tag{2.5}$$

where $A_3 = \frac{1}{2}(W_3' B_3 + B_3' W_3)$ and $G_3 = W_3 B_3^{-1}$.

Clearly, $\ell(\psi)$ is a quasi Gaussian loglikelihood in both the traditional sense that $\{v_{it}\}$ are not exactly Gaussian and the sense that $\Delta y_1$ is not exogenous but is treated as exogenous. Consequently, it may not be true that $\lim_{n\to\infty} \frac{1}{n(T-1)} S(\psi_0) \xrightarrow{p} 0$, when $T$ is fixed, where $\psi_0$ denotes the true value of the parameter vector $\psi$ (see Yang (2016) for details). Hence a necessary condition for the consistency of the CQMLEs may be violated. It is seen from the below that even $T$ increases with $n$, the CQMLEs may encounter asymptotic bias. Yang (2016) introduce a method that not only gives a consistent estimator of the model parameters when $T$ is small, but also eliminates the asymptotic bias when $T$ is large. The idea is to first find $\mathrm{E}[S(\psi_0)]$, and then adjust the quasi scores $S(\psi)$ so that the adjusted quasi score (AQS) $S^*(\psi)$ is such that $\frac{1}{\sqrt{n(T-1)}} S^*(\psi_0)$ converges to a well defined distribution.

To facilitate the discussions, denote a parametric quantity (scalar, vector or matrix) by dropping its arguments, e.g., $B_1 \equiv B_1(\lambda_1)$, $\mathbf{B}_1 \equiv \mathbf{B}_1(\lambda_1)$, $\Omega \equiv \Omega(\lambda_3)$, and denote the same quantity evaluated at the true parameter value by adding a subscript '0', e.g., $B_{10} \equiv B_1(\lambda_{10})$, $\Omega_0 \equiv \Omega(\lambda_{30})$. Let $\mathbf{C} = C \otimes I_n$ and $N = n(T-1)$. Denote $\Delta u \equiv \Delta u(\theta_0)$. The usual expectation and variance operators correspond to the true parameter values.

The following 'knowledge' about the processes in the past is necessary.

**Assumption A:** *Under Model* (2.1), *(i) the processes started $m$ periods before the start of data collection, the* 0*th period, and (ii) if $m \geq 1$, $\Delta y_0$ is independent of future errors $\{v_t, t \geq 1\}$; if $m = 0$, $y_0$ is independent of future errors $\{v_t, t \geq 1\}$.*

Under Assumption A and the assumptions that (*i*) the errors $\{v_{it}\}$ are iid across $i$ and $t$, (*ii*) the regressors are exogenous, and (*iii*) both $B_{10}^{-1}$ and $B_{30}^{-1}$ exist, Yang (2016) shows that $\mathrm{E}(\Delta Y_{-1} \Delta v') = -\sigma_{v0}^2 \mathbf{D}_{-10} \mathbf{B}_{30}^{-1}$ and $\mathrm{E}(\Delta Y \Delta v') = -\sigma_{v0}^2 \mathbf{D}_0 \mathbf{B}_{30}^{-1}$ (see also Lemma A.6, Appendix A, for more general results), which lead immediately to the followings:

$$\mathrm{E}(\Delta u' \Omega_0^{-1} \Delta Y_{-1}) = -\sigma_{v0}^2 \mathrm{tr}(\mathbf{C}^{-1} \mathbf{D}_{-10}), \tag{2.6}$$

$$\mathrm{E}(\Delta u' \Omega_0^{-1} \mathbf{W}_1 \Delta Y) = -\sigma_{v0}^2 \mathrm{tr}(\mathbf{C}^{-1} \mathbf{D}_0 \mathbf{W}_1), \tag{2.7}$$

$$\mathrm{E}(\Delta u' \Omega_0^{-1} \mathbf{W}_2 \Delta Y_{-1}) = -\sigma_{v0}^2 \mathrm{tr}(\mathbf{C}^{-1} \mathbf{D}_{-10} \mathbf{W}_2), \tag{2.8}$$

where $\mathbf{D}_{-1} \equiv \mathbf{D}_{-1}(\rho, \lambda_1, \lambda_2)$ and $\mathbf{D} \equiv \mathbf{D}(\rho, \lambda_1, \lambda_2)$, having the following expressions,

$$\mathbf{D}_{-1} = \begin{pmatrix} I_n, & 0, & \dots & 0, & 0 \\ \mathcal{B} - 2I_n, & I_n, & \dots & 0, & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathcal{B}^{T-4}(I_n - \mathcal{B})^2, & \mathcal{B}^{T-5}(I_n - \mathcal{B})^2, & \dots & \mathcal{B} - 2I_n, & I_n \end{pmatrix} \mathbf{B}_1^{-1},$$

$$\mathbf{D} = \begin{pmatrix} \mathcal{B} - 2I_n, & I_n, & \dots & 0 \\ (I_n - \mathcal{B})^2, & \mathcal{B} - 2I_n, & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{B}^{T-3}(I_n - \mathcal{B})^2, & \mathcal{B}^{T-4}(I_n - \mathcal{B})^2, & \dots & \mathcal{B} - 2I_n \end{pmatrix} \mathbf{B}_1^{-1},$$

and $\mathcal{B} \equiv \mathcal{B}(\rho, \lambda_1, \lambda_2) = B_1^{-1}(\lambda_1)(\rho I_n + \lambda_2 W_2)$.

The results (2.6)-(2.8) show that the $(\rho, \lambda_1, \lambda_2)$ elements of $\mathrm{E}[S(\psi_0)]$ are not zero. Hence, $\mathrm{plim}_{n\to\infty}\frac{1}{N}\frac{\partial}{\partial\rho}\ell_{\texttt{STLE}}(\psi_0)$, $\mathrm{plim}_{n\to\infty}\frac{1}{N}\frac{\partial}{\partial\lambda_1}\ell_{\texttt{STLE}}(\psi_0)$, and $\mathrm{plim}_{n\to\infty}\frac{1}{N}\frac{\partial}{\partial\lambda_2}\ell_{\texttt{STLE}}(\psi_0)$ are all non-zero, suggesting that the tests based on $S(\psi_0)$ obtained by treating $\Delta y_1$ as exogenous cannot be consistent in general. Thus, it is necessary to adjust the quasi scores so as to give a set of unbiased estimating functions. The adjusted quasi score (AQS) functions are:

$$S^*(\psi) = \begin{cases} \frac{1}{\sigma_v^2}\Delta X'\Omega^{-1}\Delta u(\theta), \\ \frac{1}{2\sigma_v^4}\Delta u(\theta)'\Omega^{-1}\Delta u(\theta) - \frac{N}{2\sigma_v^2}, \\ \frac{1}{\sigma_v^2}\Delta u(\theta)'\Omega^{-1}\Delta Y_{-1} + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1}), \\ \frac{1}{\sigma_v^2}\Delta u(\theta)'\Omega^{-1}\mathbf{W}_1\Delta Y + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}\mathbf{W}_1), \\ \frac{1}{\sigma_v^2}\Delta u(\theta)'\Omega^{-1}\mathbf{W}_2\Delta Y_{-1} + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1}\mathbf{W}_2), \\ \frac{1}{\sigma_v^2}\Delta u(\theta)'(C^{-1}\otimes A_3)\Delta u(\theta) - (T-1)\mathrm{tr}(G_3). \end{cases} \tag{2.9}$$

Solving $S^*(\psi) = 0$ leads to the $M$-estimator $\hat{\psi}_{\mathtt{M}}$ of $\psi$. This root-finding process can be simplified by first solving the equations for $\beta$ and $\sigma_v^2$, given $\delta = (\rho, \lambda')'$, resulting in the constrained $M$-estimators of $\beta$ and $\sigma_v^2$ as

$$\hat{\beta}(\delta) = (\Delta X'\Omega^{-1}\Delta X)^{-1}\Delta X'\Omega^{-1}(\mathbf{B}_1\Delta Y - \mathbf{B}_2\Delta Y_{-1}), \tag{2.10}$$
$$\hat{\sigma}_v^2(\delta) = \frac{1}{N}\Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta), \tag{2.11}$$

where $\Delta\hat{u}(\delta) = \Delta u(\hat{\beta}(\delta), \rho, \lambda_1, \lambda_2)$. Substituting $\hat{\beta}(\delta)$ and $\hat{\sigma}_v^2(\delta)$ into the last four components of the AQS function in (2.9) gives the concentrated AQS functions:

$$S_{\mathtt{c}}^*(\delta) = \begin{cases} \frac{1}{\hat{\sigma}_v^2(\delta)}\Delta\hat{u}(\delta)'\Omega^{-1}\Delta Y_{-1} + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1}), \\ \frac{1}{\hat{\sigma}_v^2(\delta)}\Delta\hat{u}(\delta)'\Omega^{-1}\mathbf{W}_1\Delta Y + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}\mathbf{W}_1), \\ \frac{1}{\hat{\sigma}_v^2(\delta)}\Delta\hat{u}(\delta)'\Omega^{-1}\mathbf{W}_2\Delta Y_{-1} + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1}\mathbf{W}_2), \\ \frac{1}{\hat{\sigma}_v^2(\delta)}\Delta\hat{u}(\delta)'(C^{-1}\otimes A_3)\Delta\hat{u}(\delta) - (T-1)\mathrm{tr}(G_3). \end{cases} \tag{2.12}$$

Solving the resulted concentrated estimating equations, $S_{\mathtt{c}}^*(\delta) = 0$, we obtain the unconstrained $M$-estimators $\hat{\delta}_{\mathtt{M}}$ of $\delta$. The unconstrained $M$-estimators of $\beta$ and $\sigma_v^2$ are thus $\hat{\beta}_{\mathtt{M}} \equiv \hat{\beta}(\hat{\delta}_{\mathtt{M}})$ and $\hat{\sigma}_{v,\mathtt{M}}^2 \equiv \hat{\sigma}_v^2(\hat{\delta}_{\mathtt{M}})$. Yang (2016) show that under regularity conditions the $M$-estimator $\hat{\psi}_{\mathtt{M}} = (\hat{\beta}_{\mathtt{M}}', \hat{\sigma}_{v,\mathtt{M}}^2, \hat{\delta}_{\mathtt{M}}')'$ is $\sqrt{N}$-consistent and asymptotically normal. The $M$-estimators under various constraints imposed by the various hypotheses postulated above will remain to be $\sqrt{N}$-consistent and asymptotically normal. An asymptotic result that is of particular importance to the construction of the AQS-based test statistics is that the 'normalized' $S^*(\psi_0)$ is asymptotically normal with zero mean and finite variance. It is important to note that the adjustments (2.6)-(2.8) are free from the initial conditions, and hence the resulted AQS function and the $M$-estimators are free from the initial conditions.

# 3. Adjusted Quasi Score Tests

The AQS functions given in (2.9) are the key elements in the construction of the AQS tests. In this section, we first formulate the AQS test in a unified manner, and then present details of the tests corresponding to the various joint, marginal and conditional AQS tests specified in Section 2. All the proofs are given in Appendix B.

Some general notation and convention simplify the presentations: $(i)$ $\delta$ denotes the vector of parameters in the concentrated AQS function, and $\boldsymbol{\Delta}$ the space from which $\delta$ takes values; $(ii)$ $\mathrm{tr}(\cdot)$, $|\cdot|$ and $\|\cdot\|$ denote, respectively, the trace, determinant, and Frobenius norm of a matrix; and $(iii)$ $\texttt{diag}(a_k)$ forms a diagonal matrix using the elements $\{a_k\}$ and $\texttt{blkdiag}(A_k)$ forms a block-diagonal matrix using the matrices $\{A_k\}$. The subscript '$_n$' is often dropped from an $n$-dependent quantity shall no confusion arise. $0_k$ is a $k \times 1$ vector of zeros.

## 3.1. General method

The construction of the joint and marginal AQS tests depends critically on the availability of the variance covariance (VC) matrix of the AQS function $S^*(\psi_0)$ given in (2.9), i.e., $\Gamma^*(\psi_0) = \frac{1}{N}\mathrm{Var}[S^*(\psi_0)]$. The dynamic nature of Model (2.1) makes such an estimation very difficult, as the derivation of the expression of $\Gamma^*(\psi_0)$ runs into a similar problems as the full QML estimation of the model – initial differences need to be specified or modeled when $T$ is fixed and small. To overcome this difficulty, Yang (2016) propose a martingale difference (M.D.) method, i.e., decompose the joint AQS function into a sum of M.D. sequences so that the outer-product-of-martingale-differences (OPMD) gives a consistent estimate of $\Gamma^*(\psi_0)$. As a result, the OPMD estimate of $\Gamma^*(\psi_0)$ is free from the specification of initial conditions. This together with the fact that the AQS functions are free from the specification of initial conditions lead to the AQS tests that are free from the initial conditions.

First, under Assumption A, Yang (2016) developed the following representations:

$$\Delta Y = \mathbb{R}\,\Delta\mathbf{y}_1 + \boldsymbol{\eta} + \mathbb{S}\Delta v, \quad (3.1)$$

$$\Delta Y_{-1} = \mathbb{R}_{-1}\Delta\mathbf{y}_1 + \boldsymbol{\eta}_{-1} + \mathbb{S}_{-1}\Delta v, \quad (3.2)$$

where $\mathbb{R} = \mathrm{blkdiag}(\mathcal{B}_0, \mathcal{B}_0^2, \ldots, \mathcal{B}_0^{T-1})$, $\mathbb{R}_{-1} = \mathrm{blkdiag}(I_n, \mathcal{B}_0, \ldots, \mathcal{B}_0^{T-2})$, $\boldsymbol{\eta} = \mathbb{B}\mathbf{B}_{10}^{-1}\Delta X\beta_0$, $\boldsymbol{\eta}_{-1} = \mathbb{B}_{-1}\mathbf{B}_{10}^{-1}\Delta X\beta_0$, $\mathbb{S} = \mathbb{B}\mathbf{B}_{10}^{-1}\mathbf{B}_{30}^{-1}$, $\mathbb{S}_{-1} = \mathbb{B}_{-1}\mathbf{B}_{10}^{-1}\mathbf{B}_{30}^{-1}$,

$$\mathbb{B} = \begin{pmatrix} I_n & 0 & \ldots & 0 & 0 \\ \mathcal{B}_0 & I_n & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathcal{B}_0^{T-2} & \mathcal{B}_0^{T-3} & \ldots & \mathcal{B}_0 & I_n \end{pmatrix}, \quad \textit{and} \quad \mathbb{B}_{-1} = \begin{pmatrix} 0 & 0 & \ldots & 0 & 0 \\ I_n & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathcal{B}_0^{T-3} & \mathcal{B}_0^{T-4} & \ldots & I_n & 0 \end{pmatrix}.$$

Using these representations and $\Delta u = \mathbf{B}_{30}^{-1}\Delta v$, the AQS function at $\psi_0$ is expressed as

$$
S^*(\psi_0) = \begin{cases} \Pi_1' \Delta v, \\ \Delta v' \Phi_1 \Delta v - \frac{N}{2\sigma_{v0}^2}, \\ \Delta v' \Psi_1 \Delta \mathbf{y}_1 + \Delta v' \Pi_2 + \Delta v' \Phi_2 \Delta v + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_{-10}), \\ \Delta v' \Psi_2 \Delta \mathbf{y}_1 + \Delta v' \Pi_3 + \Delta v' \Phi_3 \Delta v + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_0\mathbf{W}_1), \\ \Delta v' \Psi_3 \Delta \mathbf{y}_1 + \Delta v' \Pi_4 + \Delta v' \Phi_4 \Delta v + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_{-10}\mathbf{W}_2), \\ \Delta v' \Phi_5 \Delta v - (T-1)\mathrm{tr}(G_{30}), \end{cases} \tag{3.3}
$$

where $\Pi_1 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \Delta X$, $\Pi_2 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \boldsymbol{\eta}_{-1}$, $\Pi_3 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \mathbf{W}_1 \boldsymbol{\eta}$, $\Pi_4 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \mathbf{W}_2 \boldsymbol{\eta}_{-1}$, $\Phi_1 = \frac{1}{2\sigma_{v0}^4}(C^{-1} \otimes I_n)$, $\Phi_2 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \mathbb{S}_{-1}$, $\Phi_3 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \mathbf{W}_1 \mathbb{S}$, $\Phi_4 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \mathbf{W}_2 \mathbb{S}_{-1}$, $\Phi_5 = \frac{1}{2\sigma_{v0}^2}[C^{-1} \otimes (G_{30}' + G_{30})]$, $\Psi_1 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \mathbb{R}_{-1}$, $\Psi_2 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \mathbf{W}_1 \mathbb{R}$, $\Psi_3 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_b \mathbf{W}_2 \mathbb{R}_{-1}$, and $\mathbb{C}_b = C^{-1} \otimes B_{30}$.

The expression (3.3) is the key to the proof of the asymptotic normality of $\frac{1}{\sqrt{N}} S^*(\psi_0)$, and to the development of the OPMD estimate of the VC matrix of $S^*(\psi_0)$, so that the an AQS test can be constructed. Note that $S^*(\psi_0)$ contains three types of stochastic elements:

$$
\Pi' \Delta v, \quad \Delta v' \Phi \Delta v, \quad \text{and } \Delta v' \Psi \Delta \mathbf{y}_1,
$$

where $\Pi, \Phi$ and $\Psi$ are nonstochastic matrices (depending on $\psi_0$) with $\Pi$ being $N \times p$ or $N \times 1$, and $\Phi$ and $\Psi$ being $N \times N$. As noted in Yang (2016), the closed form expressions for variances of $\Pi' \Delta v$ and $\Delta v' \Phi \Delta v$, and their covariance can readily be derived, but the closed-form expressions for the variance of $\Delta v' \Psi \Delta \mathbf{y}_1$ and its covariances with $\Pi' \Delta v$ and $\Delta v' \Phi \Delta v$ depend on the knowledge of the distribution of $\Delta y_1$, which is unavailable. Yang (2016) went on to give a unified method of estimating the VC matrix of AQS function, the OPMD estimate, which is summarized as follows.

For a square matrix $A$, let $A^u$, $A^l$ and $A^d$ be, respectively, its upper-triangular, lower-triangular, and diagonal matrix such that $A = A^u + A^l + A^d$. Denote by $\Pi_t$, $\Phi_{ts}$ and $\Psi_{ts}$ the submatrices of $\Pi$, $\Phi$ and $\Psi$ partitioned according to $t, s = 2, \ldots, T$. Define $\Psi_{t+} = \sum_{s=2}^T \Psi_{ts}$, $t = 2, \ldots, T$, $\Theta = \Psi_{2+}(B_{30}B_{10})^{-1}$, $\Delta y_1^\circ = B_{30}B_{10}\Delta y_1$, and $\Delta y_{1t}^* = \Psi_{t+}\Delta y_1$. Define

$$
\begin{aligned}
g_{1i} &= \textstyle\sum_{t=2}^T \Pi_{it}' \Delta v_{it}, && (3.4)\\
g_{2i} &= \textstyle\sum_{t=2}^T (\Delta v_{it}\Delta \xi_{it} + \Delta v_{it}\Delta v_{it}^* - \sigma_{v0}^2 d_{it}), && (3.5)\\
g_{3i} &= \Delta v_{2i}\Delta\zeta_i + \Theta_{ii}(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2) + \textstyle\sum_{t=3}^T \Delta v_{it}\Delta y_{1it}^*, && (3.6)
\end{aligned}
$$

where for (3.5), $\xi_t = \sum_{s=2}^T (\Phi_{st}^{u\prime} + \Phi_{ts}^l)\Delta v_s$, $\Delta v_t^* = \sum_{s=2}^T \Phi_{ts}^d \Delta v_s$, and $\{d_{it}\}$ are the diagonal elements of $\mathbf{C}\Phi$; for (3.6), $\{\Delta\zeta_i\} = \Delta\zeta = (\Theta^u + \Theta^l)\Delta y_1^\circ$, and $\mathrm{diag}\{\Theta_{ii}\} = \Theta^d$. Then,

$$
\begin{aligned}
\Pi' \Delta v &= \textstyle\sum_{i=1}^n g_{1i}, && (3.7)\\
\Delta v' \Phi \Delta v - \mathrm{E}(\Delta v' \Phi \Delta v) &= \textstyle\sum_{i=1}^n g_{2i}, && (3.8)\\
\Delta v' \Psi \Delta \mathbf{y}_1 - \mathrm{E}(\Delta v' \Psi \Delta \mathbf{y}_1) &= \textstyle\sum_{i=1}^n g_{3i}, && (3.9)
\end{aligned}
$$

and $\{(g_{1i}', \ g_{2i}, \ g_{3i})', \mathcal{F}_{n,i}\}_{i=1}^n$ form a vector martingale difference (M.D.) sequence, where

$\mathcal{F}_{n,i} = \mathcal{F}_{n,0} \otimes \mathcal{G}_{n,i}$, $\{\mathcal{G}_{n,i}\}$ are the increasing sequence of $\sigma$-fields generated by $(v_{j1}, \ldots, v_{jT}, j = 1, \ldots, i), i = 1, \ldots, n$, and $\mathcal{F}_{n,0}$ is the $\sigma$-field generated by $(v_0, \Delta y_0)$.

Now, following these results, for each $\Pi_r, r = 1, 2, 3, 4$, defined in (3.3), define $g_{1ri}$ according to (3.4); for each $\Phi_r, r = 1, \ldots, 5$, defined in (3.3), define $g_{2ri}$ according to (3.5); and for each $\Psi_r, r = 1, 2, 3$, defined in (3.3), define $g_{3ri}$ according to (3.6). Define

$$\mathbf{g}_i = (g'_{11i},\ g_{21i},\ g_{31i} + g_{12i} + g_{22i},\ g_{32i} + g_{13i} + g_{23i},\ g_{33i} + g_{14i} + g_{24i},\ g_{25i})'. \tag{3.10}$$

Then, $S^*(\psi_0) = \sum_{i=1}^n \mathbf{g}_i$, and $\{\mathbf{g}_i, \mathcal{F}_{n,i}\}$ form a vector M.D. sequence. It follows that $\text{Var}[S^*(\psi_0)] = \sum_{i=1}^n \text{E}(\mathbf{g}_i \mathbf{g}_i')$. The 'average' of the outer products of the estimated $\mathbf{g}_i$, i.e.,

$$\widehat{\Gamma}^* = \tfrac{1}{N} \textstyle\sum_{i=1}^n \hat{\mathbf{g}}_i \hat{\mathbf{g}}_i', \tag{3.11}$$

thus gives a consistent estimator of the variance of $\Gamma^*(\psi_0)$, where $\hat{\mathbf{g}}_i$ is obtained by replacing $\psi_0$ in $\mathbf{g}_i$ by $\hat{\psi}_{\texttt{M}}$ and $\Delta v$ in $g_i$ by its observed counterpart $\widehat{\Delta} v$, noting that $\Delta y_1$ is observed.

Note that the AQS functions for parameter estimation given in (2.9) and the OPMD estimate of VC matrix given in (3.11) are developed based on the assumption that the errors are iid. Lemmas A.5-A.9, Appendix A, give a set of results that allow the errors to be heteroskedastic across the spatial units as in Assumption B below, which facilitate the proofs that some tests are robust against unknown cross-sectional heteroskedasticity.[2]

Now, consider the general linear hypotheses:

$$H_0 : \mathcal{C}'\psi_0 = 0,$$

where $\mathcal{C}$ is $(p+5) \times k$ with $k \leq p+4$, representing a set of linear contrasts of $\psi_0$. Let $\tilde{\mathbf{g}}_i$ be the restricted estimate of $\mathbf{g}_i$, obtained by replacing $\psi_0$ in $\mathbf{g}_i$ by $\tilde{\psi}_{\texttt{M}}$, the restricted $M$-estimator of $\psi_0$ (under $H_0$), and $v_{it}$ by $\tilde{v}_{it}$, the restricted estimates of the model errors. An AQS-based test for testing the general linear hypothesis $H_0 : \mathcal{C}'\psi_0 = 0$ is thus,

$$T_{\texttt{AQS}} = S^*(\tilde{\psi}) \big( \textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i \tilde{\mathbf{g}}_i' \big)^{-1} S^*(\tilde{\psi}), \tag{3.12}$$

Obviously, the set-up of $H_0$ above is fairly general and covers all the hypotheses postulated in Section 2. For example, one can test $H_0 : \lambda_1 = \lambda_2 = \lambda_3$ by defining the two columns of $\mathcal{C}$ as $(0'_{p+2}, 1, -1, 0)'$ and $(0'_{p+3}, 1, -1)'$. For testing $H_0^{\texttt{DP}}$: $\delta = 0$, the four columns of $\mathcal{C}$ are, respectively, $(0'_{p+1}, 1, 0'_3)'$, $(0'_{p+1}, 0, 1, 0, 0)'$, $(0'_{p+1}, 0, 0, 1, 0)'$, and $(0'_{p+1}, 0'_3, 1)'$.

The asymptotic distribution of $T_{\texttt{AQS}}$, i.e., $\chi^2_k$, can be proved under some additional regularity conditions generic to all tests, and some additional regularity conditions specific for a given test. The generic conditions are as follows.

**Assumption B:** *The innovations $v_{it}$ are such that (i) $\{v_{it}\}$ are independent across $i = 1, \ldots, n$ and $t = 0, 1, \ldots, T$ with $E(v_{it}) = 0$, (ii) $Var(v_{it}) = \sigma^2_{v0} h_{ni}$, where $0 < h_{ni} \leq c < \infty$ and $\frac{1}{n} \sum_{i=1}^n h_{ni} = 1$, and (iii) $E|v_{it}|^{4+\epsilon_0} < \infty$ for some $\epsilon_0 > 0$.*

[2]These results could potentially lead to all test statistics that are fully robust against unknown cross-sectional heteroskedasticity. However, a detailed study on this is beyond the scope of this paper.

**Assumption C:** *The time-varying regressors* $\{X_t,\ t = 0, 1, \ldots, T\}$ *are exogenous, their values are uniformly bounded, and* $\lim_{N\to\infty} \frac{1}{N}\Delta X'\Delta X$ *exists and is nonsingular.*

**Assumption D:** *(i) For* $r = 1, 2, 3$*, the elements* $w_{r,ij}$ *of* $W_r$ *are at most of order* $\iota_n^{-1}$*, uniformly in all* $i$ *and* $j$*, and* $w_{r,ii} = 0$ *for all* $i$*; (ii)* $\iota_n/n \to 0$ *as* $n \to \infty$*; (iii)* $\{W_r, r = 1, 2, 3\}$ *are uniformly bounded in both row and column sums.*

Assumption D allows the degree of spatial dependence, e.g., the number of neighbors each spatial unit has, to grow with the sample size but in a lower speed. As a result, the convergence rate of certain parameter estimators may need to be adjusted down to $\sqrt{N/\iota_n}$.[3] For certain tests where the null model contains spatial effects, Assumption B needs to be tightened by requiring $h_{ni} = 1, i = 1, \ldots, n$. Lemmas A.5-A.9 in Appendix A present more general results that could potentially lead to tests that are valid under Assumption B.

### 3.2. Joint, marginal and conditional AQS tests

To facilitate the practical applications of the AQS tests, we now present details for each of the hypothesis postulated in Section 2 so that the practitioners can pick and apply a specific test directly without going through the complicated general case. Certain tests, i.e., those specifying $\rho = 0$ at the null, can be much simpler than the general one presented above. Let $\texttt{SDPD}(\delta)$ denote the general model. An AQS test corresponds to a model reduction by setting certain elements in $\delta$ to zero, i.e., $\texttt{STPD}(\lambda)$ sets $\rho$ to 0, $\texttt{DPD}(\rho)$ sets $\lambda$ to 0, $\texttt{PD}$ sets $\delta$ to 0.

Denote a component or a subvector of the AQS vector $S^*(\psi)$ by adding a relevant subscript, e.g., $S^*_\rho(\psi)$ is the $\rho$-element and $S^*_\delta(\psi)$ is the $\delta$-subvector of $S^*(\psi)$. Denote a diagonal element or a diagonal block of $\big(\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)^{-1}$ by adding a relevant subscript, e.g., $\big(\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)^{-1}_\rho$ is the $\rho$-$\rho$ element and $\big(\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)^{-1}_\delta$ is the $\delta$-$\delta$ block of $\big(\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)^{-1}$.

**Joint test $H_0^{\texttt{PD}}$:** $\delta = 0$. Under $H_0^{\texttt{PD}}$, the model SDPD($\delta$) is reduced to the simplest PD model, and the estimation of the model at the null is simply the ordinary least squares (OLS) estimation, i.e., $\tilde\beta = (\Delta X'\mathbf{C}^{-1}\Delta X)^{-1}\Delta X'\mathbf{C}^{-1}\Delta Y$ and $\tilde\sigma_v^2 = \frac{1}{N}\Delta\tilde v'\mathbf{C}^{-1}\Delta\tilde v$, where $\Delta\tilde v = \Delta Y - \Delta X\tilde\beta$, leading to $\tilde\psi = (\tilde\beta', \tilde\sigma_v^2, 0_4')'$. Under $H_0^{\texttt{PD}}$, $B_1 = B_2 = I_n$, and $B_3 = O_n$ where $O_n$ denotes an $n \times n$ matrix of zeros. It is easy to see that $\tilde\beta$ and $\tilde\sigma_v^2$ are robust against unknown cross-sectional heteroskedasticity. Based on the fact that the $\beta$ and $\sigma_v^2$ components of the AQS vector are zero when evaluated at $\tilde\psi$, the AQS test statistics takes the form:

$$T_{\texttt{AQS}}^{\texttt{PD}} = S_\delta^{*\prime}(\tilde\psi)\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)^{-1}_\delta S_\delta^{*\prime}(\tilde\psi) \tag{3.13}$$

where $\{\tilde{\mathbf{g}}_i\}$ are obtained by evaluating $\{\mathbf{g}_i\}$ defined in (3.10) at $\psi_0 = \tilde\psi$ and $\Delta v = \Delta\tilde v$.

**Theorem 3.1.** *Under Assumptions A-D and* $H_0^{\texttt{PD}}$*, we have* $T_{\texttt{AQS}}^{\texttt{PD}} \xrightarrow{D} \chi_4^2$*. In particular, the null distribution of* $T_{\texttt{AQS}}^{\texttt{PD}}$ *is robust against unknown heteroskedasticity* $\{h_{ni}\}$*.*

[3]This typically occurs to the estimator of the spatial error parameter; see Lee (2004), Liu and Yang (2015), Su and Yang (2015), and Yang (2016). However, to simplify the proofs the asymptotic properties of the proposed tests, this feature is not explicitly reflected as the implementations of the tests do not require $\iota$.

The very attractive feature of this joint test is that it is robust against cross-sectional heteroskedasticity of unknown form as specified in Assumption B, besides being robust against nonnormality of the idiosyncratic errors $v_{it}$. The same goes to the conditional tests where under the null and given the 'condition' the model becomes pure panel model of which estimation is simply the least squares.

**Joint test $H_0^{\texttt{DPD}}$:** $\lambda = 0$. Under $H_0^{\texttt{DPD}}$, $B_1 = B_3 = I_n$, and $B_2 = \rho I_n$. The estimation of the null model goes as follows. The constrained M-estimators of $\beta$ and $\sigma_v^2$, given $\rho$, are $\tilde{\beta}(\rho) = (\Delta X'\mathbf{C}^{-1}\Delta X)^{-1}\Delta X'\mathbf{C}^{-1}(\Delta Y - \rho\Delta Y_{-1})$ and $\tilde{\sigma}_v^2(\rho) = \frac{1}{N}\Delta\tilde{v}'(\rho)\mathbf{C}^{-1}\Delta\tilde{v}(\rho)$, where $\Delta\tilde{v}(\rho) = \Delta Y - \rho\Delta Y_{-1} - \Delta X\tilde{\beta}(\rho)$. The constrained $M$-estimator of $\rho$ under $H_0^{\texttt{DPD}}$ is

$$\tilde{\rho} = \arg\left\{\tfrac{1}{\tilde{\sigma}_v^2(\rho)}\Delta\tilde{v}'(\rho)\mathbf{C}^{-1}\Delta Y_{-1} + n\big(\tfrac{1}{1-\rho} - \tfrac{1-\rho^T}{T(1-\rho)^2}\big) = 0\right\}, \tag{3.14}$$

leading to the constrained $M$ estimators of $\beta$ and $\sigma_v^2$ as $\tilde{\beta} = \tilde{\beta}(\tilde{\rho})$ and $\tilde{\sigma}_v^2 = \tilde{\sigma}_v^2(\tilde{\rho})$. The constrained $M$-estimator of $\psi$ is thus $\tilde{\psi} = \{\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho}, 0, 0, 0\}'$. The following lemma shows that the restricted $M$-estimator $\tilde{\rho}$ defined in (3.14) is robust against unknown heteroskedasticity.[4]

**Lemma 3.1.** *Under Assumptions A-D, if the parameter space $\Upsilon$ for $\rho$ is compact and $\rho_0$ is in the interior of it, the $M$-estimator $\tilde{\rho}$ for the DPD model is consistent, and so are the $M$-estimators $\tilde{\beta}$ and $\tilde{\sigma}_v^2$. Furthermore, $\sqrt{N}[(\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho})' - (\beta_0, \sigma_{v0}^2, \rho_0)'] \xrightarrow{D} N(0, \Gamma)$.*

With the $(\beta, \sigma_v^2, \rho)$-components of the AQS vector being zero when evaluated at $\tilde{\psi}$, the AQS test statistic becomes

$$T_{\texttt{AQS}}^{\texttt{DPD}} = S_\lambda^{*\prime}(\tilde{\psi})\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)_\lambda^{-1} S_\lambda^{*\prime}(\tilde{\psi}), \tag{3.15}$$

where $\{\tilde{\mathbf{g}}_i\}$ are obtained by evaluating $\{\mathbf{g}_i\}$ defined in (3.10) at $\psi_0 = \tilde{\psi}$ and $\Delta v = \Delta\tilde{v}(\tilde{\rho})$.

**Theorem 3.2.** *Under Assumptions A-D and $H_0^{\texttt{DPD}}$, we have $T_{\texttt{AQS}}^{\texttt{DPD}} \xrightarrow{D} \chi_3^2$. In particular, the asymptotic null behavior of $T_{\texttt{AQS}}^{\texttt{DPD}}$ is robust against unknown heteroskedasticity $\{h_{ni}\}$.*

Note that $\tilde{\rho}$ used in Theorem 3.2 needs not be the constrained $M$-estimator defined in (3.14), and can be replaced by any $\sqrt{N}$-consistent and heteroskedasticity robust estimator.

**Marginal test $H_0^{\texttt{STPD}}$:** $\rho = 0$. Under the null, $B_2 = \lambda_2 W_2$. The constrained M-estimator $\tilde{\lambda}$ of $\lambda$ solves the following estimating equations:

$$\begin{cases} \frac{1}{\tilde{\sigma}_v^2(\lambda)}\Delta\tilde{u}(\lambda)'\Omega^{-1}\mathbf{W}_1\Delta Y + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}\mathbf{W}_1) = 0, \\ \frac{1}{\tilde{\sigma}_v^2(\lambda)}\Delta\tilde{u}(\lambda)'\Omega^{-1}\mathbf{W}_2\Delta Y_{-1} + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1}\mathbf{W}_2) = 0, \\ \frac{1}{\tilde{\sigma}_v^2(\lambda)}\Delta\tilde{u}(\lambda)'(C^{-1} \otimes A_3)\Delta\tilde{u}(\lambda) - (T-1)\mathrm{tr}(G_3) = 0, \end{cases}$$

[4]The concentrated AQS function for $\rho$ contained in (3.14) clearly shows that the $M$-estimator is not only consistent when $T$ is fixed but also eliminates the bias of order $O(T^{-1})$. In contrast, the estimator based on the unadjusted score is inconsistent when $T$ is fixed and has a bias of order $O(T^{-1})$ when $T$ grows with $n$. See Hahn and Kuersteiner (2002), and Yang (2016) for more discussions and related works on the DPD model.

where $\Delta\tilde{u}(\lambda) = \mathbf{B}_1\Delta Y - \lambda_2\mathbf{W}_2\Delta Y_{-1} - \Delta X\tilde{\beta}(\lambda)$, and $\tilde{\beta}(\lambda)$ and $\tilde{\sigma}_v^2(\lambda)$ are obtained from (2.10) and (2.11) by setting $\rho = 0$. Let $\tilde{\beta} = \tilde{\beta}(\tilde{\lambda})$, $\tilde{\sigma}_v^2 = \tilde{\sigma}_v^2(\tilde{\lambda})$, and $\tilde{\psi} = \{\tilde{\beta}', \tilde{\sigma}_v^2, 0, \tilde{\lambda}'\}'$. The AQS test for $H_0^{\texttt{STPD}}$ has the form

$$T_{\texttt{AQS}}^{\texttt{STPD}} = S_\rho^*(\tilde{\psi})\big[\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)_\rho^{-1}\big]^{\frac{1}{2}}, \tag{3.16}$$

where $\{\tilde{\mathbf{g}}_i\}$ are obtained by evaluating $\{\mathbf{g}_i\}$ defined in (3.10) at $\psi_0 = \tilde{\psi}$ and $\Delta v = \tilde{\mathbf{B}}_3^{-1}\tilde{u}(\tilde{\lambda})$.

**Theorem 3.3.** *Under Assumptions A-D and $H_0^{\texttt{STPD}}$, if further, (i) $h_{ni} = 1, i = 1, \ldots, n$, (ii) $B_{10}^{-1}$ and $B_{30}^{-1}$ are uniformly bounded in both row and column sums and (iii) $\tilde{\lambda}$ is $\sqrt{N}$-consistent, then we have $T_{\texttt{AQS}}^{\texttt{STPD}} \xrightarrow{D} N(0,1)$.*

The asymptotic normality of the test statistic $T_{\texttt{AQS}}^{\texttt{STPD}}$ at the null typically requires that the constrained estimator $\tilde{\lambda}$ be $\sqrt{N}$-consistent. This is implied by the general result of Yang (2015) under homoskedastic errors and hence is not discussed in this paper. Furthermore, $\tilde{\lambda}$ needs not be the constrained $M$-estimator discussed above, and any other estimator this is $\sqrt{N}$-consistent can be used. See the proof of the theorem given in Appendix B. The result of Theorem 3.3 shows that the asymptotic null behavior of the test statistic $T_{\texttt{AQS}}^{\texttt{STPD}}$ is not fully robust against unknown heteroskedasticity $\{h_{ni}\}$. Use of the results in Lemma A.9 may provide a version of the variance estimator that is robust against unknown heteroskedasticity but it is not clear how to make the AQS function and $\tilde{\lambda}$ also robust against unknown heteroskedasticity. However, our Monte Carlo results show that this test and its improved version given in Sec. 4 are quite robust against unknown heteroskedasticity. These discussions apply to all the tests given below, as well as their improved versions presented in Sec. 4.

**Marginal test $H_0^{\texttt{SDPDr}}$:** $\lambda_r = 0$, where $r$ can be 1, or 2 or 3, giving three marginal tests corresponding one specific type of spatial effects. Among these three marginal tests, the test of $H_0^{\texttt{SDPD2}}$: $\lambda_2 = 0$ is the most interesting one as under $H_0^{\texttt{SDPD2}}$ the model is reduced to the popular SDPD model with `SL` and `SE` (or `SARAR`) effects. We consider only this case as the others can be handled in the similar manner. Under $H_0^{\texttt{SDPD2}}$, $B_2 = \rho I_n$. The constrained M-estimators $(\tilde{\rho}, \tilde{\lambda}_1, \tilde{\lambda}_3)$ of $(\rho, \lambda_1, \lambda_3)$ solve the following estimating equations:

$$\begin{cases} \frac{1}{\tilde{\sigma}_v^2(\rho,\lambda_1,\lambda_3)}\Delta\tilde{u}(\rho,\lambda_1,\lambda_3)'\Omega^{-1}\Delta Y_{-1} + \text{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1}) = 0, \\ \frac{1}{\tilde{\sigma}_v^2(\rho,\lambda_1,\lambda_3)}\Delta\tilde{u}(\rho,\lambda_1,\lambda_3)'\Omega^{-1}\mathbf{W}_1\Delta Y + \text{tr}(\mathbf{C}^{-1}\mathbf{D}\mathbf{W}_1) = 0, \\ \frac{1}{\tilde{\sigma}_v^2(\rho,\lambda_1,\lambda_3)}\Delta\tilde{u}(\rho,\lambda_1,\lambda_3)'(C^{-1}\otimes A_3)\Delta\tilde{u}(\rho,\lambda_1,\lambda_3) - (T-1)\text{tr}(G_3) = 0, \end{cases}$$

where $\Delta\tilde{u}(\rho,\lambda_1,\lambda_3) = \mathbf{B}_1\Delta Y - \rho\Delta Y_{-1} - \Delta X\tilde{\beta}(\rho,\lambda_1,\lambda_3)$, and $\tilde{\beta}(\rho,\lambda_1,\lambda_3)$ and $\tilde{\sigma}_v^2(\rho,\lambda_1,\lambda_3)$ are obtained from (2.10) and (2.11) by setting $\lambda_2 = 0$. Let $\tilde{\beta} = \tilde{\beta}(\tilde{\rho},\tilde{\lambda}_1,\tilde{\lambda}_3)$, $\tilde{\sigma}_v^2 = \tilde{\sigma}_v^2(\tilde{\rho},\tilde{\lambda}_1,\tilde{\lambda}_3)$, and $\tilde{\psi} = \{\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho}, \tilde{\lambda}_1, 0, \tilde{\lambda}_3\}'$. The AQS test for $H_0^{\texttt{SDPD2}}$ has the form

$$T_{\texttt{AQS}}^{\texttt{SPDD2}} = S_{\lambda_2}^*(\tilde{\psi})\big[\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)_{\lambda_2}^{-1}\big]^{\frac{1}{2}}, \tag{3.17}$$

where $\{\tilde{\mathbf{g}}_i\}$ are $\{\mathbf{g}_i\}$ defined in (3.10), evaluated at $\psi_0 = \tilde{\psi}$ and $\Delta v = \tilde{\mathbf{B}}_3^{-1}\Delta\tilde{u}(\tilde{\rho},\tilde{\lambda}_1,\tilde{\lambda}_3)$.

**Theorem 3.4.** *Under Assumptions A-D and $H_0^{\texttt{SDPD2}}$, if further, (i) $h_{ni} = 1, i = 1, \ldots, n$, (ii) $B_{10}^{-1}$ and $B_{30}^{-1}$ are uniformly bounded in both row and column sums and (iii) $(\tilde{\rho}, \tilde{\lambda}_1, \tilde{\lambda}_3)$ are $\sqrt{N}$-consistent, then we have $T_{\texttt{AQS}}^{\texttt{SDPD2}} \xrightarrow{D} N(0,1)$.*

**Joint test $H_0^{\texttt{SDPD4}}$:** $\lambda_1 = \lambda_2 = 0$. This is an interesting joint test as under the null the model reduces to a popular SDPD model with spatial error only, which was studied by Su and Yang (2015) under fixed $T$ and with initial observations modeled. Under $H_0^{\texttt{SDPD4}}$, $B_1 = I_n$ and $B_2 = \rho I_n$. The $M$-estimators $\tilde{\rho}$ and $\tilde{\lambda}_3$ of $\rho$ and $\lambda_3$ solve the estimating equations:

$$\begin{cases} \frac{1}{\tilde{\sigma}_v^2(\rho,\lambda_3)} \Delta\tilde{u}(\rho,\lambda_3)'\Omega^{-1}\Delta Y_{-1} + \text{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1}) = 0, \\ \frac{1}{\tilde{\sigma}_v^2(\rho,\lambda_3)} \Delta\tilde{u}(\rho,\lambda_3)'(C^{-1} \otimes A_3)\Delta\tilde{u}(\rho,\lambda_3) - (T-1)\text{tr}(G_3) = 0, \end{cases}$$

where $\Delta\tilde{u}(\rho,\lambda_3) = \Delta Y - \rho\Delta Y_{-1} - \Delta X\tilde{\beta}(\rho,\lambda_1,\lambda_3)$, and $\tilde{\beta}(\rho,\lambda_3)$ and $\tilde{\sigma}_v^2(\rho,\lambda_3)$ are obtained from (2.10) and (2.11) by setting $\lambda_1 = \lambda_2 = 0$. Let $\tilde{\beta} = \tilde{\beta}(\tilde{\rho},\tilde{\lambda}_3)$, $\tilde{\sigma}_v^2 = \tilde{\sigma}_v^2(\tilde{\rho},\tilde{\lambda}_3)$, and $\tilde{\psi} = \{\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho}, 0, 0, \tilde{\lambda}_3\}'$. The AQS test for $H_0^{\texttt{SDPD4}}$ has the form:

$$T_{\texttt{AQS}}^{\texttt{DPD}} = S_{\rho,\lambda_3}^{*\prime}(\tilde{\psi})\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)_{\rho,\lambda_3}^{-1} S_{\rho,\lambda_3}^{*\prime}(\tilde{\psi}), \tag{3.18}$$

where $\{\tilde{\mathbf{g}}_i\}$ are $\{\mathbf{g}_i\}$ defined in (3.10), evaluated at $\psi_0 = \tilde{\psi}$ and $\Delta v = \tilde{\mathbf{B}}_3^{-1}\Delta\tilde{u}(\tilde{\rho},\tilde{\lambda}_3)$.

**Theorem 3.5.** *Under Assumptions A-D and $H_0^{\texttt{SDPD4}}$, if further, (i) $h_{ni} = 1, i = 1, \ldots, n$, (ii) $B_{30}^{-1}$ is uniformly bounded in both row and column sums and (iii) $\tilde{\rho}$ and $\tilde{\lambda}_3$ are $\sqrt{N}$-consistent, then we have $T_{\texttt{AQS}}^{\texttt{SDPD4}} \xrightarrow{D} \chi_2^2$.*

**Joint test $H_0^{\texttt{SDPD5}}$:** $\lambda_2 = \lambda_3 = 0$. Under the null hypothesis, the model reduces to another popular model, the SDPD model with only the spatial lag effect. Under $H_0^{\texttt{SDPD5}}$, $B_2 = \rho I_n$ and $B_3 = I_n$. The constrained $M$-estimators $\tilde{\rho}$ and $\tilde{\lambda}_1$ of $\rho$ and $\lambda_1$ solve the following equations:

$$\begin{cases} \frac{1}{\tilde{\sigma}_v^2(\rho,\lambda_1)} \Delta\tilde{v}(\rho,\lambda_1)'\Omega^{-1}\Delta Y_{-1} + \text{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1}) = 0, \\ \frac{1}{\tilde{\sigma}_v^2(\rho,\lambda_1)} \Delta\tilde{v}(\rho,\lambda_1)'\Omega^{-1}\mathbf{W}_1\Delta Y + \text{tr}(\mathbf{C}^{-1}\mathbf{D}\mathbf{W}_1) = 0, \end{cases}$$

where $\Delta\tilde{v}(\rho,\lambda_1) = \mathbf{B}_1\Delta Y - \rho\Delta Y_{-1} - \Delta X\tilde{\beta}(\rho,\lambda_1)$, and $\tilde{\beta}(\rho,\lambda_1)$ and $\tilde{\sigma}_v^2(\rho,\lambda_1 = \lambda_3)$ are obtained from (2.10) and (2.11) by setting $\lambda_2 = \lambda_3 = 0$. Let $\tilde{\beta} = \tilde{\beta}(\tilde{\rho},\tilde{\lambda}_1)$, $\tilde{\sigma}_v^2 = \tilde{\sigma}_v^2(\tilde{\rho},\tilde{\lambda}_1)$, and $\tilde{\psi} = \{\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho}, \tilde{\lambda}_1, 0, 0\}'$. The AQS test for $H_0^{\texttt{SDPD5}}$ has the form

$$T_{\texttt{AQS}}^{\texttt{DPD}} = S_{\rho,\lambda_1}^{*\prime}(\tilde{\psi})\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)_{\rho,\lambda_1}^{-1} S_{\rho,\lambda_1}^{*\prime}(\tilde{\psi}), \tag{3.19}$$

where $\{\tilde{\mathbf{g}}_i\}$ are $\{\mathbf{g}_i\}$ defined in (3.10), evaluated at $\psi_0 = \tilde{\psi}$ and $\Delta v = \Delta\tilde{v}(\tilde{\rho},\tilde{\lambda}_1)$.

**Theorem 3.6.** *Under Assumptions A-D and $H_0^{\texttt{SDPD5}}$, if further, (i) $h_{ni} = 1, i = 1, \ldots, n$, (ii) $B_{10}^{-1}$ is uniformly bounded in both row and column sums and (iii) $\tilde{\rho}$ and $\tilde{\lambda}_1$ are $\sqrt{N}$-consistent, then we have $T_{\texttt{AQS}}^{\texttt{SDPD5}} \xrightarrow{D} \chi_2^2$.*

**Joint test $H_0^{\texttt{SPD}}$:** $\rho = \lambda_2 = 0$. Under the null, $B_2 = 0$ and $\mathbf{D} = -\mathbf{C}\mathbf{B}_1^{-1}$. The model becomes the static SARAR model. The constrained M-estimators $\tilde{\lambda}_1$ and $\tilde{\lambda}_3$ of $\lambda_1$ and $\lambda_3$ solve the following estimating equations (see also Lee and Yu (2010)):

$$\begin{cases} \frac{1}{\tilde{\sigma}_v^2(\lambda_1,\lambda_3)}\Delta\tilde{u}(\lambda_1,\lambda_3)'\Omega^{-1}\mathbf{W}_1\Delta Y - (T-1)\text{tr}(B_1^{-1}W_1) = 0, \\ \frac{1}{\tilde{\sigma}_v^2(\lambda_1,\lambda_3)}\Delta\tilde{u}(\lambda_1,\lambda_3)'(C^{-1}\otimes A_3)\Delta\tilde{u}(\lambda_1,\lambda_3) - (T-1)\text{tr}(G_3) = 0, \end{cases}$$

where $\Delta\tilde{u}(\lambda_1,\lambda_3) = \mathbf{B}_1\Delta Y - \Delta X\tilde{\beta}(\lambda_1,\lambda_3)$, and $\tilde{\beta}(\lambda_1,\lambda_3)$ and $\tilde{\sigma}_v^2(\lambda_1,\lambda_3)$ are obtained from (2.10) and (2.11) by setting $\rho = \lambda_2 = 0$. Let $\tilde{\beta} = \tilde{\beta}(\tilde{\lambda}_1,\tilde{\lambda}_3)$, $\tilde{\sigma}_v^2 = \tilde{\sigma}_v^2(\tilde{\lambda}_1,\tilde{\lambda}_3)$, and $\tilde{\psi} = \{\tilde{\beta}', \tilde{\sigma}_v^2, 0, \tilde{\lambda}_1, 0, \tilde{\lambda}_3\}'$. The AQS test for $H_0^{\texttt{SPD}}$ has the form

$$T_{\texttt{AQS}}^{\texttt{SPD}} = S_{\rho,\lambda_2}^{*}(\tilde{\psi})'\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i'\big)_{\rho,\lambda_2}^{-1} S_{\rho,\lambda_2}^{*}(\tilde{\psi}), \tag{3.20}$$

where $\{\tilde{\mathbf{g}}_i\}$ are $\{\mathbf{g}_i\}$ defined in (3.10), but evaluated at $\psi_0 = \tilde{\psi}$ and $\Delta v = \tilde{\mathbf{B}}_3^{-1}\tilde{u}(\tilde{\lambda}_1,\tilde{\lambda}_3)$.

**Theorem 3.7.** *Under Assumptions A-D and $H_0^{\texttt{SPD}}$, if further, (i) $h_{ni} = 1, i = 1,\ldots,n$, (ii) $B_{10}^{-1}$ and $B_{30}^{-1}$ are uniformly bounded in both row and column sums and (iii) $\tilde{\lambda}_1$ and $\tilde{\lambda}_3$ are $\sqrt{N}$-consistent, then we have $T_{\texttt{AQS}}^{\texttt{SPD}} \xrightarrow{D} \chi_2^2$.*

Conditional tests are those for testing whether the model can be further reduced, given that it has already been reduced. For example, $H_{c0}^{\texttt{PD1}}$: $\lambda_1 = 0$, given $\lambda_2 = \lambda_3 = 0$; $H_{c0}^{\texttt{PD3}}$: $\lambda_3 = 0$, given $\lambda_1 = \lambda_2 = 0$; $H_{c0}^{\texttt{STPD}}$: $\rho = 0$, given $\lambda_2 = 0$. The last conditional test says that based on the model without $\lambda_2$, we want to see if $\rho = 0$, i.e., if the model SDPD$(\rho,\lambda_1,\lambda_3)$ can be reduced to SPD$(\lambda_1,\lambda_3)$, a static spatial panel data model. The conditional tests conditional upon $\rho = 0$ are the tests of model reduction for the SDPD$(\lambda)$ model, and the LM-type of tests have been developed by, e.g., Debarsy and Erther (2010) and Baltagi and Yang (2013a) for models with homoskedastic models, and Born and Breitung (2011) and Baltagi and Yang (2013b) for models with heteroskedastic errors. All these conditional tests can be easily developed based on the general methodology presented above. Some conditional tests are robust against heteroskedasticity in light of Theorems 3.1 and 3.2, and some can be made to be robust against heteroskedasticity in light of Baltagi and Yang (2013b) and Lemmas A.8 and A.9. Given the fact that the OPMD estimator of the VC matrix of estimating equations are robust against cross-sectional heteroskedasticity or can be made so in light of Lemma A.9, any AQS or SAQS test can be made to be heteroskedasticity robust, provided the AQS function is robust. However, it is not clear how to adjust the AQS functions so that they are fully robust against heteroskedasticity due the presence of initial differences $\Delta y_1$. All the tests developed above can be implemented in a unified manner based on the general expressions of the AQS function given in (2.9) or (3.3), and the general OPMD estimate of its VC matrix given in (3.11). For each specific test, all it is necessary is to change the definitions of the matrices $B_r, r = 1,2,3$ according to the null hypothesis, and modify the user-supplied function that does root-finding. Matlab codes are available from the author upon request.

# 4. Finite Sample Improved AQS Tests

The joint and marginal AQS tests presented above are simple but may not be satisfactory when $n$ is not large enough. The reason is that the variability from the estimation of $\beta$ and $\sigma_v^2$ are not taken into account when constructing the test statistics. It is thus desirable to find ways to improve the finite sample performance of these tests. Clearly, after $\beta_0$ and $\sigma_v^2$ being replaced by $\hat{\beta}(\delta_0)$ and $\hat{\sigma}_v(\delta_0)$ in the last four elements of $S^*(\psi_0)$ given in (2.9), the concentrated AQS functions no longer have mean zero, although they do asymptotically. Furthermore, the variance of the concentrated AQS functions may also be affected. Thus, re-adjustments on the mean and variance may help improving the finite sample performance of the AQS tests (see Baltagi and Yang 2013a,b).

## 4.1. General method

Rewrite the concentrated AQS function given in (2.12) as

$$S_{\mathtt{c}}^*(\delta) = \begin{cases} \frac{1}{\hat{\sigma}_v^2(\delta)}\big[\Delta\hat{u}(\delta)'\Omega^{-1}\Delta Y_{-1} + \phi_1\Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta)\big], \\ \frac{1}{\hat{\sigma}_v^2(\delta)}\big[\Delta\hat{u}(\delta)'\Omega^{-1}\mathbf{W}_1\Delta Y + \phi_2\Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta)\big], \\ \frac{1}{\hat{\sigma}_v^2(\delta)}\big[\Delta\hat{u}(\delta)'\Omega^{-1}\mathbf{W}_2\Delta Y_{-1} + \phi_3\Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta)\big], \\ \frac{1}{\hat{\sigma}_v^2(\delta)}\big[\Delta\hat{u}(\delta)'(C^{-1}\otimes A_3)\Delta\hat{u}(\delta) - \phi_4\Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta)\big], \end{cases} \tag{4.1}$$

where $\phi_1 = \frac{1}{N}\text{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1})$, $\phi_2 = \frac{1}{N}\text{tr}(\mathbf{C}^{-1}\mathbf{D}\mathbf{W}_1)$, $\phi_3 = \frac{1}{N}\text{tr}(\mathbf{C}^{-1}\mathbf{D}_{-1}\mathbf{W}_2)$ and $\phi_4 = \frac{1}{n}\text{tr}(G_3)$, and consider the numerator $S_{\mathtt{c,N}}^*(\delta)$ of $S_{\mathtt{c}}^*(\delta)$, i.e., the vector without the scale multiplier $\frac{1}{\hat{\sigma}_v^2(\delta)}$. The ideas are: finding the mean of $S_{\mathtt{c,N}}^*(\delta_0)$ and recentering, and then finding the variance estimate of the recentered $S_{\mathtt{c,N}}^*(\delta_0)$ and restandardizing.

Letting $\Omega^{\frac{1}{2}}$ be the symmetric square root matrix of $\Omega$, and $\Delta X^* = \Omega^{-\frac{1}{2}}\Delta X$, we have

$$\Omega^{-\frac{1}{2}}\Delta\hat{u}(\delta) = \mathbf{M}\Omega^{-\frac{1}{2}}(\mathbf{B}_1\Delta Y - \mathbf{B}_2\Delta Y_{-1}),$$

where $\mathbf{M} = I_N - \Delta X^*(\Delta X^{*\prime}\Delta X^*)^{-1}\Delta X^{*\prime}$ is a projection matrix. Noting that $\mathbf{M}\Delta X^* = 0$, and that at the true parameter values $\Omega_0^{-\frac{1}{2}}(\mathbf{B}_{10}\Delta Y - \mathbf{B}_{20}\Delta Y_{-1}) = \Delta X^*\beta_0 + \Omega_0^{-\frac{1}{2}}\mathbf{B}_{30}^{-1}\Delta v$, the numerator $S_{\mathtt{c,N}}^*(\delta_0)$ of the concentrated AQS functions can be further written as

$$S_{\mathtt{c,N}}^*(\delta_0) = \begin{cases} \Delta v'\mathbf{B}_{30}'^{-1}\mathbf{M}_0^*\Delta Y_{-1} + \phi_{10}\Delta v'\mathbf{M}_0^{**}\Delta v, \\ \Delta v'\mathbf{B}_{30}'^{-1}\mathbf{M}_0^*\mathbf{W}_1\Delta Y + \phi_{20}\Delta v'\mathbf{M}_0^{**}\Delta v, \\ \Delta v'\mathbf{B}_{30}'^{-1}\mathbf{M}_0^*\mathbf{W}_2\Delta Y_{-1} + \phi_{30}\Delta v'\mathbf{M}_0^{**}\Delta v, \\ \Delta v'\mathbf{M}_0^{**}(C\otimes G_{30}^{\circ})\mathbf{M}_0^{**}\Delta v - \phi_{40}\Delta v'\mathbf{M}_0^{**}\Delta v, \end{cases} \tag{4.2}$$

where $G_3^{\circ} = \frac{1}{2}(G_3' + G_3)$, $\mathbf{M}^* = \Omega^{-\frac{1}{2}}\mathbf{M}\Omega^{-\frac{1}{2}}$ and $\mathbf{M}^{**} = \mathbf{B}_3'^{-1}\mathbf{M}^*\mathbf{B}_3^{-1}$.[5]

[5]Noting that $\mathbf{M}^* = \Omega^{-1} - \Omega^{-1}\Delta X(\Delta X'\Omega^{-1}\Delta X)^{-1}\Delta X'\Omega^{-1}$, the calculations of $\Omega^{\frac{1}{2}}$ and its inverse, which can be computationally demanding when $N$ is large, are avoided in real applications.

By Lemma A.6, it is easy to see that $\mathrm{E}[S^{*}_{\mathtt{c,N}}(\delta_0)] = (\mu_{\rho_0}, \mu_{\lambda_{10}}, \mu_{\lambda_{20}}, \mu_{\lambda_{30}})'$, where $\mu_{\rho_0} = \sigma^2_{v0}\mathrm{tr}[(\mathbf{B}'_{30}\mathbf{B}_{30})^{-1}\mathbf{M}^{*}_0(\phi_{10}\mathbf{C} - \mathbf{D}_{-10})]$, $\mu_{\lambda_{10}} = \sigma^2_{v0}\mathrm{tr}[(\mathbf{B}'_{30}\mathbf{B}_{30})^{-1}\mathbf{M}^{*}_0(\phi_{20}\mathbf{C} - \mathbf{W}_1\mathbf{D}_0)]$, $\mu_{\lambda_{20}} = \sigma^2_{v0}\mathrm{tr}[(\mathbf{B}'_{30}\mathbf{B}_{30})^{-1}\mathbf{M}^{*}_0(\phi_{30}\mathbf{C} - \mathbf{W}_2\mathbf{D}_{-10})]$, and $\mu_{\lambda_{30}} = \sigma^2_{v0}\mathrm{tr}[\mathbf{M}^{**}_0(C \otimes G^{\circ}_{30} - \phi_{40}\mathbf{C})]$, giving the recentered AQS function as:

$$S^{**}_{\mathtt{c,N}}(\delta_0) = S^{*}_{\mathtt{c,N}}(\delta_0) - (\mu_{\rho_0}, \mu_{\lambda_{10}}, \mu_{\lambda_{20}}, \mu_{\lambda_{30}})'. \tag{4.3}$$

To develop an OPMD estimate of the VC matrix of $S^{**}_{\mathtt{c,N}}(\delta_0)$, we have by (3.1) and (3.2),

$$S^{**}_{\mathtt{c,N}}(\delta_0) = \begin{cases} \Delta v'\Psi_1\Delta\mathbf{y}_1 + \Delta v'\Pi_1 + \Delta v'\Phi_1\Delta v - \mu_{\rho_0}, \\ \Delta v'\Psi_2\Delta\mathbf{y}_1 + \Delta v'\Pi_2 + \Delta v'\Phi_2\Delta v - \mu_{\lambda_{10}}, \\ \Delta v'\Psi_3\Delta\mathbf{y}_1 + \Delta v'\Pi_3 + \Delta v'\Phi_3\Delta v - \mu_{\lambda_{20}}, \\ \Delta v'\Phi_4\Delta v - \mu_{\lambda_{30}}, \end{cases} \tag{4.4}$$

where $\Pi_1 = \mathbf{B}'^{-1}_{30}\mathbf{M}^{*}_0\boldsymbol{\eta}_{-1}$, $\Pi_2 = \mathbf{B}'^{-1}_{30}\mathbf{M}^{*}_0\mathbf{W}_1\boldsymbol{\eta}$, $\Pi_3 = \mathbf{B}'^{-1}_{30}\mathbf{M}^{*}_0\mathbf{W}_2\boldsymbol{\eta}_{-1}$; $\Phi_1 = \mathbf{B}'^{-1}_{30}\mathbf{M}^{*}_0\mathbb{S}_{-1} + \phi_{10}\mathbf{M}^{**}_0$, $\Phi_2 = \mathbf{B}'^{-1}_{30}\mathbf{M}^{*}_0\mathbf{W}_1\mathbb{S} + \phi_{20}\mathbf{M}^{**}_0$, $\Phi_3 = \mathbf{B}'^{-1}_{30}\mathbf{M}^{*}_0\mathbf{W}_2\mathbb{S}_{-1} + \phi_{30}\mathbf{M}^{**}_0$, $\Phi_4 = \mathbf{M}^{**}_0(C \otimes G^{\circ}_{30})\mathbf{M}^{**}_0 - \phi_{40}\mathbf{M}^{**}_0$; $\Psi_1 = \mathbf{B}'^{-1}_{30}\mathbf{M}^{*}_0\mathbb{R}_{-1}$, $\Psi_2 = \mathbf{B}'^{-1}_{30}\mathbf{M}^{*}_0\mathbf{W}_1\mathbb{R}$, $\Psi_3 = \mathbf{B}'^{-1}_{30}\mathbf{M}^{*}_0\mathbf{W}_2\mathbb{R}_{-1}$.

Similar to $\{\mathbf{g}_i\}$ defined based on (3.3), we define $\{\mathbf{g}^{\circ}_i\}$ based on (4.4). Now, $\{\mathbf{g}^{\circ}_i\}$ are functions of unknown parameters $\delta_0$ and the unobserved errors $\Delta v$. Replacing $\delta_0$ by $\tilde{\delta}$ and $\Delta v$ by $\Delta\tilde{v}(\tilde{\delta})$ in $\{\mathbf{g}^{\circ}_i\}$ to give $\{\tilde{\mathbf{g}}^{\circ}_i\}$, one obtains an OPMD estimate of $\frac{1}{N}\mathrm{Var}[S^{**}_{\mathtt{c,N}}(\delta_0)]$ as

$$\widehat{\Gamma}^{**} = \frac{1}{N}\sum_{i=1}^{n}\tilde{\mathbf{g}}^{\circ}_i\tilde{\mathbf{g}}^{\circ\prime}_i, \tag{4.5}$$

and the standardized AQS (SAQS) test statistic for testing $H_0 : \mathcal{C}'\delta_0 = 0$ as

$$T_{\mathtt{SAQS}} = S^{**}_{\mathtt{c,N}}(\tilde{\delta})\big(\textstyle\sum_{i=1}^{n}\tilde{\mathbf{g}}^{\circ}_i\tilde{\mathbf{g}}^{\circ\prime}_i\big)^{-1}S^{**}_{\mathtt{c,N}}(\tilde{\delta}), \tag{4.6}$$

where $\mathcal{C}$ is $4 \times k$ ($k \leq 4$), and the limiting null distribution of $T_{\mathtt{SAQS}}$ can be shown to be $\chi^2_k$. Monte Carlo results presented in the following section show that the SAQS tests offer much improvements over the AQS tests when $n$ is not large.

### 4.2. Improved joint, marginal and conditional AQS tests

Again, to facilitate the practical applications of the standardized AQS tests, we present details of the tests for each of the hypothesis postulated in Section 2. Clearly, all the standardized AQS tests can be expressed in a single form as in (4.6). However, each test has a certain specific property and hence deserves a detailed study. Similarly, one can formulate the tests in the same manner as the AQS tests presented above by using only the standardized AQS component(s) of the parameter(s) that the test concerns, although the remaining components evaluated at the restricted estimates of $\delta$ and $\Delta v$ are no longer strictly zero (but negligible). Similarly, denote the component or subvector of $S^*_{\mathtt{c,N}}(\delta)$, and the element or block of $\big(\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)^{-1}$ by adding a relevant subscript, e.g., $S^{**}_{\mathtt{c,N},\lambda}(\delta)$ is the $\lambda$-component of $S^{**}_{\mathtt{c,N}}(\delta)$, and $\big(\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)^{-1}_\lambda$ is the $\lambda$-$\lambda$ block of $\big(\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)^{-1}$. The details for each tests are as follows.

**Joint test $H_0^{\mathtt{PD}}$:** $\delta = 0$. Under $H_0^{\mathtt{PD}}$, the model SDPD($\delta$) is reduced to the simplest PD model, and the estimation of the model at the null is simply the ordinary least square (OLS) estimation. The standardized AQS test for testing $H_0^{\mathtt{PD}}$: $\delta = 0$ is thus,

$$T^{\mathtt{PD}}_{\mathtt{SAQS}} = S^{**\prime}_{\mathtt{c,N}}(0)\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)^{-1} S^{**}_{\mathtt{c,N}}(0), \tag{4.7}$$

where $\{\tilde{\mathbf{g}}_i^\circ\}$ are obtained by evaluating $\mathbf{g}_i^\circ$ at $\delta_0 = 0$ and $\Delta v = \Delta Y - \Delta X\tilde{\beta}(0)$.

**Corollary 4.1.** *Under the assumptions of Theorem 3.1 and $H_0^{\mathtt{PD}}$, $T^{\mathtt{PD}}_{\mathtt{SAQS}} \xrightarrow{D} \chi^2_4$. Thus, the asymptotic null behavior of $T^{\mathtt{PD}}_{\mathtt{SAQS}}$ is robust against unknown heteroskedasticity $\{h_{ni}\}$.*

**Joint test $H_0^{\mathtt{DPD}}$:** $\lambda = 0$. Let $\tilde{\delta} = (\tilde{\rho}, 0, 0, 0)'$ and $\Delta\tilde{v} = \Delta Y - \tilde{\rho}\Delta Y_{-1} - \Delta X\tilde{\beta}(\tilde{\delta})$. The standardized AQS test takes either the form of (4.6), or the following simpler form,

$$T^{\mathtt{DPD}}_{\mathtt{SAQS}} = S^{**\prime}_{\mathtt{c,N},\lambda}(\tilde{\delta})\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)^{-1}_\lambda S^{**}_{\mathtt{c,N},\lambda}(\tilde{\delta}), \tag{4.8}$$

where $\{\tilde{\mathbf{g}}_i^\circ\}$ are $\{\mathbf{g}_i^\circ\}$ with $\delta_0$ and $\Delta v$ replaced by $\tilde{\delta}$ and $\Delta\tilde{v}$.

**Corollary 4.2.** *Under the assumptions of Theorem 3.2 and $H_0^{\mathtt{DPD}}$, $T^{\mathtt{DPD}}_{\mathtt{SAQS}} \xrightarrow{D} \chi^2_3$. Thus, the asymptotic null behavior of $T^{\mathtt{DPD}}_{\mathtt{SAQS}}$ is robust against unknown heteroskedasticity $\{h_{ni}\}$.*

Strictly speaking, the test statistic defined in (4.8) is not identical to the corresponding one obtained from (4.6). This is because the constrained $M$-estimator $\tilde{\rho}$ solves the concentrated AQS function for $\rho$ as in (3.14), and thus the $\rho$-element of $S^{**}_{\mathtt{c,N}}(\tilde{\delta})$ is not identically 0. However, such a difference is negligible. The same issue applies to the tests given below.

**Marginal test $H_0^{\mathtt{STPD}}$:** $\rho = 0$. Let $\tilde{\delta} = (0, \tilde{\lambda}')'$ and $\Delta\tilde{v} = \tilde{\mathbf{B}}_3^{-1}[\tilde{\mathbf{B}}_1\Delta Y - \tilde{\lambda}_2\mathbf{W}_2 Y_{-1} - \Delta X\tilde{\beta}(\tilde{\delta})]$. The standardized AQS test takes either the form of (4.6), or the simpler form,

$$T^{\mathtt{STPD}}_{\mathtt{SAQS}} = S^{**}_{\mathtt{c,N},\rho}(\tilde{\delta})\big[\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)^{-1}_\rho\big]^{\frac{1}{2}}, \tag{4.9}$$

where $\{\tilde{\mathbf{g}}_i^\circ\}$ are $\{\mathbf{g}_i^\circ\}$ with $\delta_0$ and $\Delta v$ replaced by $\tilde{\delta}$ and $\Delta\tilde{v}$.

**Corollary 4.3.** *Under the assumptions of Theorem 3.3 and $H_0^{\mathtt{STPD}}$, $T^{\mathtt{STPD}}_{\mathtt{SAQS}} \xrightarrow{D} N(0,1)$.*

**Marginal test $H_0^{\mathtt{SDPD2}}$:** $\lambda_2 = 0$. Let $\tilde{\delta} = (\tilde{\rho}, \tilde{\lambda}_1, 0, \tilde{\lambda}_3)'$ and $\Delta\tilde{v} = \tilde{\mathbf{B}}_3^{-1}[\tilde{\mathbf{B}}_1\Delta Y - \mathbf{B}_2 Y_{-1} - \Delta X\tilde{\beta}(\tilde{\delta})]$. The standardized AQS test takes either the form of (4.6), or the simpler form,

$$T^{\mathtt{SDPD2}}_{\mathtt{SAQS}} = S^{**}_{\mathtt{c,N},\lambda_2}(\tilde{\delta})\big[\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)^{-1}_{\lambda_2}\big]^{\frac{1}{2}}, \tag{4.10}$$

where $\{\tilde{\mathbf{g}}_i^\circ\}$ are $\{\mathbf{g}_i^\circ\}$ with $\delta_0$ and $\Delta v$ replaced by $\tilde{\delta}$ and $\Delta\tilde{v}$. Similarly, the test $T^{\circ\mathtt{SDPD1}}_{\mathtt{AQS}}$

for testing $H_0^{\texttt{SDPD1}}$: $\lambda_1 = 0$, and the test $T_{\texttt{AQS}}^{\circ\texttt{SDPD3}}$ for testing $H_0^{\texttt{SDPD3}}$: $\lambda_3 = 0$ are developed. However, these tests are less interested and hence details are not given.

**Corollary 4.4.** *Under the assumptions of Theorem 3.4 and* $H_0^{\texttt{SDPD2}}$, $T_{\texttt{SAQS}}^{\texttt{SDPD2}} \xrightarrow{D} N(0,1)$.

**Joint test $H_0^{\texttt{SDPD4}}$:** $\lambda_1 = \lambda_2 = 0$. Let $\tilde{\delta} = (\tilde{\rho}, 0, 0, \tilde{\lambda}_3)'$ and $\Delta\tilde{v} = \tilde{\mathbf{B}}_3^{-1}[\Delta Y - \tilde{\rho}\Delta Y_{-1} - \Delta X\tilde{\beta}(\tilde{\delta})]$. The SAQS test takes either the form of (4.6), or the following simpler form,

$$T_{\texttt{SAQS}}^{\texttt{SDPD4}} = S_{\texttt{c,N},\lambda_1,\lambda_3}^{**\prime}(\tilde{\delta})\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)_{\lambda_1,\lambda_3}^{-1} S_{\texttt{c,N},\lambda_1,\lambda_2}^{**}(\tilde{\delta}), \tag{4.11}$$

where $\{\tilde{\mathbf{g}}_i^\circ\}$ are $\{\mathbf{g}_i^\circ\}$ with $\delta_0$ and $\Delta v$ replaced by $\tilde{\delta}$ and $\Delta\tilde{v}$.

**Corollary 4.5.** *Under the assumptions of Theorem 3.5 and* $H_0^{\texttt{SDPD4}}$, $T_{\texttt{SAQS}}^{\texttt{SDPD4}} \xrightarrow{D} \chi_2^2$.

**Joint test $H_0^{\texttt{SDPD5}}$:** $\lambda_2 = \lambda_3 = 0$. Let $\tilde{\delta} = (\tilde{\rho}, \tilde{\lambda}_1, 0, 0)'$ and $\Delta\tilde{v} = \Delta Y - \mathbf{B}_2\Delta Y_{-1} - \Delta X\tilde{\beta}(\tilde{\delta})$. The SAQS test takes either the form of (4.6), or the following simpler form,

$$T_{\texttt{SAQS}}^{\texttt{SDPD5}} = S_{\texttt{c,N},\lambda_2,\lambda_3}^{**\prime}(\tilde{\delta})\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)_{\lambda_2,\lambda_3}^{-1} S_{\texttt{c,N},\lambda_2,\lambda_3}^{**}(\tilde{\delta}), \tag{4.12}$$

where $\{\tilde{\mathbf{g}}_i^\circ\}$ are $\{\mathbf{g}_i^\circ\}$ with $\delta_0$ and $\Delta v$ replaced by $\tilde{\delta}$ and $\Delta\tilde{v}$.

**Corollary 4.6.** *Under the assumptions of Theorem 3.6 and* $H_0^{\texttt{SDPD5}}$, $T_{\texttt{SAQS}}^{\texttt{SDPD5}} \xrightarrow{D} \chi_2^2$.

**Joint test $H_0^{\texttt{SPD}}$:** $\rho = \lambda_2 = 0$. Let $\tilde{\delta} = (0, \tilde{\lambda}_1, 0, \tilde{\lambda}_3)'$ and $\Delta\tilde{v} = \tilde{\mathbf{B}}_3^{-1}[\tilde{\mathbf{B}}_1\Delta Y - \Delta X\tilde{\beta}(\tilde{\delta})]$. The SAQS test takes either the form of (4.6), or the simpler form,

$$T_{\texttt{SAQS}}^{\texttt{SPD}} = S_{\texttt{c,N},\rho,\lambda_2}^{**}(\tilde{\delta})'\big(\textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i^\circ \tilde{\mathbf{g}}_i^{\circ\prime}\big)_{\rho,\lambda_2}^{-1} S_{\texttt{c,N},\rho,\lambda_2}^{**}(\tilde{\delta}), \tag{4.13}$$

where $\{\tilde{\mathbf{g}}_i^\circ\}$ are $\{\mathbf{g}_i^\circ\}$ with $\delta_0$ and $\Delta v$ replaced by $\tilde{\delta}$ and $\Delta\tilde{v}$.

**Corollary 4.7.** *Under the assumptions of Theorem 3.7 and* $H_0^{\texttt{SPD}}$, $T_{\texttt{SAQS}}^{\texttt{SPD}} \xrightarrow{D} \chi_2^2$.

All the conditional AQS tests discussed in Section 3 have their counterparts based on the standardized AQS function. Similar to the case of the regular AQS tests presented in Sec. 3, the standardized AQS tests can also be implemented in a unified manner based on the general expressions (4.3) or (4.4), and the VC matrix estimate defined in (4.5). Matalb codes are available from the author upon request.

## 5. Monte Carlo Simulation

Extensive Monte Carlo experiments are carried out to investigate the finite sample performance of the proposed AQS test and standardized AQS (SAQS) test, in terms of the size of the tests, the means and standard deviations (sds) of the test statistics at the null hypothesis.

The following data generating process (DGP) is followed:

$$y_t = \rho y_{t-1} + \lambda_1 W_1 y_t + \lambda_2 W_2 y_{t-1} + \beta_0 \iota_n + X_t\beta_1 + Z\gamma + \mu + u_t, \;\; u_t = \lambda_3 W_3 u_t + v_t,$$

with certain parameter(s) being dropped corresponding to each specific test, for generating observations at the null. The elements of $X_t$ are generated as in Yang (2016), and the elements of $Z$ are randomly generated from $Bernoulli(0.5)$. The spatial weight matrices are generated according to Rook, Queen contiguity, or group interaction schemes.[6] The values of $(\beta_0, \beta_1, \gamma, \sigma_\mu, \sigma_v)$ are set to $(5, 1, 1, 1, 1)$, $T = 3$ or 6, and $n = (50, 100, 200, 500)$. Each set of Monte Carlo results is based on 5000 samples. The error ($v_t$) distributions can be ($i$) normal, ($ii$) normal mixture (10% $N(0, 4)$ and 90% $N(0, 1)$), or ($iii$) lognormal.[7] The fixed effects $\mu$ are generated according to either $\frac{1}{T}\sum_{t=1}^{T} X_t + e$, where $e \sim (0, I_N)$, resulting in the fixed effects that are correlated with the regressors.

We only report results, in the form of empirical means, standard deviations (sds), and the sizes at the nominal levels $10\%, 5\%$ and $1\%$, corresponding to the seven pairs of tests described in details in Sections 3 and 4 with $T = 3$. The results for other tests lead to similar conclusions and are available from the author upon request. Furthermore, Monte Carlo experiments are repeated with $T = 6$ and the results (not reported for brevity but available upon request) show similar patterns as the case of $T = 3$.

Table 1 presents results for testing $H_0^{\texttt{PD}}$: $\delta = 0$. When $n$ is not large, the AQS test can be severely oversized, whereas the standardized AQS test can be slightly undersized. As $n$ increases, the mean, sd, and size of the SAQS test quickly approach to their nominal values corresponding to the $\chi_4^2$ distribution, but even when $n = 500$, the AQS test shows a clear departure from $\chi_4^2$ with its mean, sd and size significantly larger than the nominal values. As shown in Theorem 3.1 and Corollary 4.1, these pair of tests are robust against cross-sectional heteroskedasticity. The results given in the last panel of Table 1 confirm this. The results further reveal that heteroskedasticity deteriorates the finite sample performance of AQS test, but not the finite sample performance of the SAQS test.

Tables 2a and 2b presents results for testing $H_0^{\texttt{DPD}}$: $\lambda = 0$, allowing $\rho$ to be present in the model as a free parameter, with errors being homoskedastic or heteroskedastic, respectively. The results show an excellent performance of the standardized AQS test with its empirical means, sds and sizes being very close to their nominal values even when $n = 50$. In contrast, the regular AQS test may have sever size distortions when $n$ is not so large, which converges to its nominal level in a significantly slower speed than that of the standardized AQS test. The true value of $\rho$ does not have a significant effect on both tests. As shown by Theorem 3.2 and Corollary 4.2, both AQS and SAQS tests are robust against unknown heteroskedasticity.

[6]The Rook and Queen schemes are standard. For group interaction, we first generate $k = \sqrt{n}$ groups of sizes $n_g \sim U(.5\bar{n}, 1.5\bar{n})$, $g = 1, \cdots, k$ and $\bar{n} = n/k$, and then adjust $n_g$ so that $\sum_{g=1}^{k} n_g = n$. See Yang (2016) for details in generating these spatial layouts.

[7]In both (ii) and (iii), the generated errors are standardized to have mean zero and variance $\sigma_v^2$.

The results reported in Table 2b confirm these theoretical results. From the results, we also observe that the unknown heteroskedasticity seems affect the finite sample performance of the AQS test more than that of the SAQS test.

Table 3 presents results for testing $H_0^{\mathtt{SDPD4}}$: $\lambda_1 = \lambda_2 = 0$, allowing $\rho$ and $\lambda_3$ to be present in the model as free parameters. The null model corresponds to an interesting model popular in practical applications, showing the importance of such a test. The results again show an excellent performance of the standardized AQS test, which significantly improves the regular AQS tests. As $n$ increases, the null distributions of both tests converge to $\chi_2^2$ quite fast.

Table 4 presents results for another interesting test $H_0^{\mathtt{SDPD5}}$ where the null specifies $\lambda_2 = \lambda_3 = 0$, allowing $\rho$ and $\lambda_1$ to be present in the model as free parameters. The results again show that the AQS test can be severely oversized, while the standardized AQS test exhibits a moderate size distortion when $n$ is not large which quickly disappears as $n$ increases. In this case, the true values of $\rho$ and $\lambda_3$ seem to have a noticeable effect on the performance of both tests. Both tests are consistent in the sense that as $n$ increases their null distributions converge to the standard normal distribution.

Table 5 presents results for testing $H_0^{\mathtt{SDPD2}}$: $\lambda_2 = 0$, allowing $\rho$, $\lambda_1$ and $\lambda_3$ to be present in the model as free parameters. The results again show that both tests are consistent in the sense that their sizes, means, and sds at the null converge to the nominal values, but the standardized AQS test significantly improves the regular AQS test in finite samples. It is interesting to note that in this testing situation the error distribution plays a more significant role on the finite sample performance of the tests with a skewed error distribution (DGP 3) leading to a more severe sized distortion.

Table 6 presents results for another interesting test $H_0^{\mathtt{STPD}}$: $\rho = 0$, allowing $\lambda$ to be present in the model as free parameters. The results show again that the standardized AQS test significantly improves regular the AQS test in terms of sizes, and means and sds of the test statistics in finite samples, but both tests are consistent in the sense that as $n$ increase, the size, mean and sd of the test statistics at the null converge to their nominal values. The true values of the $\lambda's$ do not seem to have a noticeable impact on the finite sample performance of the tests. Finally, Table 7 presents results for testing $H_0^{\mathtt{SPD}}$: $\rho = \lambda_2 = 0$, treating $\lambda_1$ and $\lambda_3$ as free parameters. The results exhibit similar patterns as the results for the other tests.

Some additional Monte Carlo experiments are run and the results are not reported for brevity. The results for the conditional tests reveal similar patterns. The results under a larger $T = 6$ show that the finite sample performance of both tests improve, but the general observations remain. We have also run extensive Monte Carlo experiments to check the robustness of those tests which in theory are not fully robust against unknown heteroskedasticity. The results (not reported for brevity) show that these tests are quite robust against mild departure from homoskedasticity of the errors. Considering the fact that the standardized AQS tests are as simple to implement as the regular AQS tests, it is recommended the standardized AQS tests be used in the practical applications, unless $n$ is fairly large.

## 6. Conclusions and Discussions

General methods for constructing tests for the existence/nonexistence of dynamic and/or spatial effects in the fixed effects panel data model are introduced, based on the adjust quasi scores and their martingale difference representations. The methods for standardizing the tests for improved finite sample performance are also introduced. The standardized versions of the tests are shown to be as simple as the non-standardized versions but are more reliable in finite samples, hence are recommended for the empirical applications. The results presented in the paper show that the general methodology for constructing tests are promising. While certain tests are fully robust against unknown cross-sectional heteroskedasticity, the others are not although Monte Carlo simulation has demonstrated that they are also quite robust against mild departure from homoskedasticity. It is interesting to develop tests that are fully robust against unknown cross-sectional heteroskedasticity, and the general theoretical results presented in Lemmas A.5-A.9, Appendix A, offer a possibility. However, there are two difficulties: (*i*) these tests typically involve the estimation of submodels and the consistent estimators of the model parameters under heteroskedasticity may not be available, and (*ii*) the way to further adjust the AQS functions to be heteroskedasticity robust is not clear given the presence of the initial differences in the AQS functions. Thus, a detailed study of this issue is beyond the scope of this paper and will be carried out in a future research.

## Appendix A: Some Fundamental Results

The development and the proofs of theoretical results reported in this paper depend critically on the following lemmas.

**Lemma A.1.** (Kelejian and Prucha, 1999; Lee, 2002). *Let $\{A_n\}$ and $\{B_n\}$ be two sequences of $n \times n$ matrices that are uniformly bounded in both row and column sums. Let $C_n$ be a sequence of conformable matrices whose elements are uniformly $O(\iota_n^{-1})$. Then*

*(i) the sequence $\{A_n B_n\}$ are uniformly bounded in both row and column sums,*

*(ii) the elements of $A_n$ are uniformly bounded and $tr(A_n) = O(n)$, and*

*(iii) the elements of $A_n C_n$ and $C_n A_n$ are uniformly $O(\iota_n^{-1})$.*

**Lemma A.2.** (Lee, 2004, p.1918). *For $W_r$ and $B_r$, $r = 1, 2$, defined in Model (2.1), if $\|W_r\|$ and $\|B_{r0}^{-1}\|$ are uniformly bounded, where $\|\cdot\|$ is a matrix norm, then $\|B_r^{-1}\|$ is uniformly bounded in a neighborhood of $\lambda_{r0}$.*

**Lemma A.3.** (Lee, 2004, p.1918). *Let $X_n$ be an $n \times p$ matrix. If the elements $X_n$ are uniformly bounded and $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ exists and is nonsingular, then $P_n = X_n(X_n'X_n)^{-1}X_n'$ and $M_n = I_n - P_n$ are uniformly bounded in both row and column sums.*

**Lemma A.4.** *Let $\{A_n\}$ be a sequence of $n \times n$ matrices that are uniformly bounded in either row or column sums. Suppose that the elements $a_{n,ij}$ of $A_n$ are $O(\iota_n^{-1})$ uniformly in all $i$ and $j$. Let $v_n$ be a random $n$-vector satisfying Assumption B, and $b_n$ a random $n$-vector independent of $v_n$ such that $\{\mathrm{E}(b_{ni}^2)\}$ are of uniform order $O(\iota_n^{-1})$. Then,*

*(i)* $\mathrm{E}(v_n' A_n v_n) = O(\frac{n}{\iota_n})$, *(ii)* $\mathrm{Var}(v_n' A_n v_n) = O(\frac{n}{\iota_n})$,

*(iii)* $\mathrm{Var}(v_n' A_n v_n + b_n' v_n) = O(\frac{n}{\iota_n})$, *(iv)* $v_n' A_n v_n = O_p(\frac{n}{\iota_n})$,

*(v)* $v_n' A_n v_n - \mathrm{E}(v_n' A_n v_n) = O_p((\frac{n}{\iota_n})^{\frac{1}{2}})$, *(vi)* $v_n' A_n b_n = O_p((\frac{n}{\iota_n})^{\frac{1}{2}})$.

**Proof of Lemma A.4.** Simply modify the proof of Lemma A.5 of Yang (2016) by allowing heteroskedasticity $\{h_{ni}\}$ in $v_n$. ■

**Lemma A.5.** *Let $\{\Phi_n\}$ be a sequence of $n \times n$ matrices with row and column sums uniformly bounded, and elements $\phi_{n,ij}$ of uniform order $O(\iota_n^{-1})$. Let $v_n$ be an $n \times 1$ random vector satisfying Assumption B. Let $b_n = \{b_{ni}\}$ be a sequence of $n \times 1$ random vectors, independent of $v_n$, such that (i) $\{\mathrm{E}(b_{ni}^2)\}$ are of uniform order $O(\iota_n^{-1})$, (ii) $sup_i E|b_{ni}|^{2+\epsilon_0} < \infty$, and (iii) $\frac{\iota_n}{n}\sum_{i=1}^n h_{ni}(b_{ni}^2 - \mathrm{E}b_{ni}^2) = o_p(1)$. Let $\mathcal{H}_n = \mathrm{diag}\{h_{ni}, i = 1, \ldots, n\}$ and define the bilinear-quadratic form:*

$$Q_n = b_n' v_n + v_n' \Phi_n v_n - \sigma_v^2 tr(\Phi_n \mathcal{H}_n),$$

*and let $\sigma_{Q_n}^2$ be the variance of $Q_n$. If $lim_{n\to\infty} \iota_n^{1+2/\epsilon_0}/n = 0$ and $\{\frac{\iota_n}{n}\sigma_{Q_n}^2\}$ are bounded away from zero, then $Q_n/\sigma_{Q_n} \xrightarrow{d} N(0,1)$.*

**Proof of Lemma A.5.** This lemma extends Lemma A.5 of Yang (2016) by allowing heteroskedasticity on $\{v_i\}$, and thus the proof proceeds similarly. Assume (W.L.O.G.) $\Phi_n$ is symmetric with elements $\phi_{n,ij}$. Write $Q_n = \sum_{i=1}^n [b_{ni}v_i + v_i\xi_{ni} + \phi_{n,ii}(v_i^2 - \sigma_v^2 h_{ni})] \equiv \sum_{i=1}^n \mathsf{Y}_{ni}$, where $\xi_{ni} = 2\sum_{j=1}^{i-1} \phi_{n,ij}v_j$. Let $\mathcal{G}_i = \sigma(v_1, \ldots, v_i)$ be the $\sigma$-fields generated by $(v_1, \ldots, v_i), i = 1, \ldots, n$, and $\mathcal{F}_{n0}$ the $\sigma$-field generated by $b_n$. By independence between $b_n$ and $v_n$, $\mathcal{F}_{ni} = \mathcal{F}_{n0} \times \mathcal{G}_i$ is the $\sigma$-field generated by $(b_n, v_1, \ldots, v_i)$. By construction, $\mathcal{F}_{n,i-1} \subseteq \mathcal{F}_{ni}$. Clearly, $\mathsf{Y}_{ni}$ is $\mathcal{F}_{ni}$-measurable and $\xi_{ni}$ is $\mathcal{F}_{n,i-1}$-measurable. It follows that $\mathrm{E}(\mathsf{Y}_{ni}|\mathcal{F}_{n,i-1}) = b_{ni}\mathrm{E}(v_i) + \mathrm{E}(v_i)\xi_{ni} + \phi_{n,ii}\mathrm{E}(v_i^2 - \sigma_{v0}^2 h_{ni}) = 0$, and hence $\{\mathsf{Y}_{ni}, \mathcal{F}_{ni},\ 1 \le i \le n\}$ form a M.D. array, and $\sigma_{Q_n}^2 = \sum_{i=1}^n E(Y_{ni}^2)$. Define $\mathsf{Z}_{ni} = \mathsf{Y}_{ni}/\sigma_{Q_n}$. Then, $\{\mathsf{Z}_{ni}, \mathcal{F}_{ni},\ 1 \le i \le n\}$ also form a M.D. array. The proof of the lemma thus amounts to verify the conditions for the central limit theorem (CLT) for M.D. arrays, e.g., the condition (A.1) or (A.3) and condition (A.2) of Theorem A.1 in Kelejian and Prucha (2001):

$$\begin{array}{ll} (a) & \sum_{i=1}^n \mathrm{E}[\mathrm{E}(|\mathsf{Z}_{ni}|^{2+\delta}|\mathcal{F}_{n,i-1})] \longrightarrow 0, \quad \text{for some } \delta > 0; \\ (b) & \sum_{i=1}^n \mathrm{E}(\mathsf{Z}_{ni}^2|\mathcal{F}_{n,i-1}) \xrightarrow{p} 1. \end{array}$$

The detail for the proof of $(a)$ follows closely that of Theorem 1 of Kelejian and Prucha (2001), where the quantities $|b_{ni}|$, $b_{ni}^2$, $|b_{ni}|^q$ are replaced by their expectations, and with reference to the proof of Lemma A.13 of Lee (2004b) to take care of unbounded $\iota_n$.

To prove $(b)$, we have $\sum_{i=1}^n \mathrm{E}[\mathsf{Z}_{ni}^2|\mathcal{F}_{n,i-1}] - 1 = \sigma_{Q_n}^{-2}\sum_{i=1}^n[\mathrm{E}(\mathsf{Y}_{ni}^2|\mathcal{F}_{n,i-1}) - \mathrm{E}(\mathsf{Y}_{ni}^2)]$. With $\mathsf{Y}_{ni} = b_{ni}v_i + v_i\xi_{ni} + \phi_{n,ii}(v_i^2 - \sigma_v^2 h_{ni})$, it is easy to see that

$$\begin{aligned} \tfrac{\iota_n}{n}\sum_{i=1}^n[\mathrm{E}(\mathsf{Y}_{ni}^2|\mathcal{F}_{n,i-1}) - \mathrm{E}(\mathsf{Y}_{ni}^2)] &= \sigma_{v0}^2\tfrac{\iota_n}{n}\sum_{i=1}^n h_{ni}(b_{ni}^2 - \mathrm{E}b_{ni}^2) + \sigma_{v0}^2\tfrac{\iota_n}{n}\sum_{i=1}^n h_{ni}(\xi_{ni}^2 - \tau_{ni}^2) \\ &+ 2\sigma_{v0}^2\tfrac{\iota_n}{n}\sum_{i=1}^n b_{ni}\xi_{ni} + 2\tfrac{\iota_n}{n}\sum_{i=1}^n \phi_{n,ii}\mu_{3i}\xi_{ni} + 2\tfrac{\iota_n}{n}\sum_{i=1}^n \phi_{n,ii}\mu_{3i}(b_{ni} - \mathrm{E}b_{ni}) \equiv \sum_{r=1}^5 Q_r, \end{aligned}$$

where $\tau_{ni}^2 = \mathrm{Var}(\xi_{ni}) = 4\sigma_{v0}^2\sum_{j=1}^{i-1}\phi_{n,ij}^2 h_{nj}$ and $\mu_{3i} = \mathrm{E}(v_{it}^3)$. Thus, to show (b) it is sufficient to show that $Q_r \xrightarrow{p} 0, r = 1, \ldots, 5$.

First the result for $Q_1$ follows from the assumption (iii) of Lemma A.5, and the result for $Q_5$ follows from the assumption (ii) of Lemma A.5 and Chebyshev inequality. Now,

$$Q_2 = \sigma_{v0}^2\tfrac{\iota_n}{n}\sum_{i=1}^n h_{ni}(\xi_{ni}^2 - \tau_{ni}^2) = 4\sigma_{v0}^2\tfrac{\iota_n}{n}\sum_{j=1}^{n-1} a_{nj}(v_j^2 - \sigma_{v0}^2 h_{ni}) + 8\sigma_{v0}^2\tfrac{\iota_n}{n}\sum_{j=1}^{n-1} v_j\varepsilon_{nj}.$$

where $a_{nj} = \sum_{i=j+1}^n h_{ni}\phi_{n,ij}^2$, $\varepsilon_{nj} = \sum_{k=1}^{j-1} c_{n,ik}v_k$, and $c_{n,ik} = \sum_{i=j+1}^n h_{ni}\phi_{n,ij}\phi_{n,ik}$. Clearly, both $\{(v_j^2 - \sigma_{v0}^2 h_{ni}),\ \mathcal{G}_i\}$ and $\{v_j\varepsilon_{nj},\ \mathcal{G}_i\}$ are M.D. arrays, and hence their convergence in probability to zero is proved by applying the weak law of large numbers (WLLN) for M.D. arrays of Davidson (1994, p. 299). It follows that $Q_2 \xrightarrow{p} 0$.

By applying Chebyshev inequality, we show that $Q_3 \xrightarrow{p} 0$. Now, it is easy to see that $Q_4 = 2\frac{\iota_n}{n}\sum_{j=1}^{n-1} d_{n,j}v_j$ where $d_{n,j} = \sum_{i=j+1}^n \phi_{n,ii}\mu_{3i}\phi_{ij}$. Thus, the convergence of $Q_4$ is proved by applying Chebyshev inequality, noting that both $h_{ni}$ and $\phi_{n,ii}$ are uniformly bounded. ■

**Lemma A.6.** *Suppose Assumption A and Assumptions B(i)-(ii) hold for Model* (2.1).

*Assume further that (i) the time-varying regressors $X_t$ are exogenous, and (ii) both $B_{10}^{-1}$ and $B_{30}^{-1}$ exist. Then, we have,*

$$\begin{aligned} \mathrm{E}(\Delta Y_{-1}\Delta v') &= -\sigma_{v0}^2 \mathbf{D}_{-10}\mathbf{B}_{30}^{-1}\mathbf{H}, & \text{(A.1)} \\ \mathrm{E}(\Delta Y \Delta v') &= -\sigma_{v0}^2 \mathbf{D}_0 \mathbf{B}_{30}^{-1}\mathbf{H}, & \text{(A.2)} \end{aligned}$$

*where $\mathbf{D}_{-1} \equiv \mathbf{D}_{-1}(\rho, \lambda_1, \lambda_2)$ and $\mathbf{D} \equiv \mathbf{D}(\rho, \lambda_1, \lambda_2)$ are given below* (2.8)*, $\mathbf{H} = I_{T-1} \otimes \mathcal{H}_n$, and $\mathcal{H}_n = \mathrm{diag}\{h_{ni}, i = \ldots, n\}$.*

**Proof of Lemma A.6.** From (2.2), we have $\Delta y_t = \mathcal{B}_0 \Delta y_{t-1} + B_{10}^{-1}\Delta X_t + B_{10}^{-1}B_{30}^{-1}\Delta v_t,\ \ t = 2, \ldots, T$. Under Assumption A, if $m \geq 1$, then

$$\mathrm{E}(\Delta y_1 \Delta v_2') = B_{10}^{-1}B_{30}^{-1}\mathrm{E}(\Delta v_1 \Delta v_2') = -\sigma_{v0}^2 B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n;$$

if $m = 0$, then $\mathrm{E}(\Delta y_1 \Delta v_2') = B_{10}^{-1}B_{30}^{-1}\mathrm{E}(y_1 \Delta v_2') = B_{10}^{-1}B_{30}^{-1}\mathrm{E}(v_1 \Delta v_2') = -\sigma_{v0}^2 B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n$.

The above result remains to be true for $t \geq 2$, i.e.,

$$\mathrm{E}(\Delta y_t \Delta v_{t+1}') = B_{10}^{-1}B_{30}^{-1}\mathrm{E}(v_t \Delta v_{t+1}') = -\sigma_{v0}^2 B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n.$$

It follows that, for $t \geq 2$,

$$\begin{aligned} \mathrm{E}(\Delta y_t \Delta v_t') &= \mathcal{B}_0 \mathrm{E}(\Delta y_{t-1}\Delta v_t') + B_{10}^{-1}B_{30}^{-1}\mathrm{E}(\Delta v_t \Delta v_t') \\ &= -\sigma_{v0}^2 \mathcal{B}_0 B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n + 2\sigma_{v0}^2 B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n \\ &= \sigma_{v0}^2(2I_n - \mathcal{B}_0)B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n; \\ \mathrm{E}(\Delta y_{t+1}\Delta v_t') &= \mathcal{B}_0 \mathrm{E}(\Delta y_t \Delta v_t') + B_{10}^{-1}B_{30}^{-1}\mathrm{E}(\Delta v_{t+1}\Delta v_t') \\ &= \sigma_{v0}^2(2\mathcal{B}_0 - \mathcal{B}_0^2)B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n - \sigma_{v0}^2 B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n \\ &= -\sigma_{v0}^2(I_n - \mathcal{B}_0)^2 B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n. \end{aligned}$$

Finally, for $t \geq 3$, we have, $\mathrm{E}(\Delta y_t \Delta v_2') = -\sigma_{v0}^2 \mathcal{B}_0^{t-3}(I_n - \mathcal{B}_0)^2 B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n$, for $t \geq 4$, we have, $\mathrm{E}(\Delta y_t \Delta v_3') = -\sigma_{v0}^2 \mathcal{B}_0^{t-4}(I_n - \mathcal{B}_0)^2 B_{10}^{-1}B_{30}^{-1}\mathcal{H}_n$, and so forth. Summarize these, we obtain the results of Lemma (A.6). ■

**Lemma A.7.** *Under the assumptions of Lemma A.6, we have*

$$\mathrm{E}[S^*(\psi_0)] = \begin{cases} 0_{p+1}, \\ \mathrm{tr}(\mathbf{D}_{-10}\mathbf{C}^{-1}) - \mathrm{tr}[\mathbf{D}_{-10}(C^{-1} \otimes (B_{30}^{-1}\mathcal{H}_n B_{30}))], \\ \mathrm{tr}(\mathbf{W}_1 \mathbf{D}_0 \mathbf{C}^{-1}) - \mathrm{tr}[\mathbf{W}_1 \mathbf{D}_0 (C^{-1} \otimes (B_{30}^{-1}\mathcal{H}_n B_{30}))], \\ \mathrm{tr}(\mathbf{W}_2 \mathbf{D}_{-10}\mathbf{C}^{-1}) - \mathrm{tr}[\mathbf{W}_2 \mathbf{D}_{-10}(C^{-1} \otimes (B_{30}^{-1}\mathcal{H}_n B_{30}))], \\ (T-1)[\mathrm{tr}(G_{30}\mathcal{H}) - \mathrm{tr}(G_{30})]. \end{cases}$$

**Proof of Lemma A.7.** It is immediate from the results of Lemma A.6. ■

The following discussions and results extend those in Section 3.1 by allowing unknown heteroskedasticity $\{h_{ni}\}$ in $v_{it}$ across $i$. They may potentially lead to test statistics that are fully robust against unknown heteroskedasticity. Let $S^{\circ}(\psi_0)$ be the centered version of $S^*(\psi_0)$

according to Lemma A.6, i.e.,

$$S^{\circ}(\psi_0) = \begin{cases} \frac{1}{\sigma_{v0}^2}\Delta X'\Omega_0^{-1}\Delta u(\theta_0), \\ \frac{1}{2\sigma_{v0}^4}\Delta u(\theta_0)'\Omega_0^{-1}\Delta u(\theta_0) - \frac{N}{2\sigma_{v0}^2}, \\ \frac{1}{\sigma_{v0}^2}\Delta u(\theta_0)'\Omega_0^{-1}\Delta Y_{-1} + \mathrm{tr}[\mathbf{D}_{-10}(C^{-1}\otimes(B_{30}^{-1}\mathcal{H}_n B_{30}))], \\ \frac{1}{\sigma_{v0}^2}\Delta u(\theta_0)'\Omega_0^{-1}\mathbf{W}_1\Delta Y + \mathrm{tr}[\mathbf{W}_1\mathbf{D}_0(C^{-1}\otimes(B_{30}^{-1}\mathcal{H}_n B_{30}))], \\ \frac{1}{\sigma_{v0}^2}\Delta u(\theta_0)'\Omega_0^{-1}\mathbf{W}_2\Delta Y_{-1} + \mathrm{tr}[\mathbf{W}_2\mathbf{D}_{-10}(C^{-1}\otimes(B_{30}^{-1}\mathcal{H}_n B_{30}))], \\ \frac{1}{\sigma_{v0}^2}\Delta u(\theta_0)'(C^{-1}\otimes A_{30})\Delta u(\theta_0) - (T-1)\mathrm{tr}(G_{30}\mathcal{H}_n), \end{cases}$$

which can be further written by (3.1) and (3.2) as

$$S^{\circ}(\psi_0) = \begin{cases} \Pi_1'\Delta v, \\ \Delta v'\Phi_1\Delta v - \frac{N}{2\sigma_{v0}^2}, \\ \Delta v'\Psi_1\Delta\mathbf{y}_1 + \Delta v'\Pi_2 + \Delta v'\Phi_2\Delta v + \mathrm{tr}[\mathbf{D}_{-10}(C^{-1}\otimes(B_{30}^{-1}\mathcal{H}_n B_{30}))], \\ \Delta v'\Psi_2\Delta\mathbf{y}_1 + \Delta v'\Pi_3 + \Delta v'\Phi_3\Delta v + \mathrm{tr}[\mathbf{W}_1\mathbf{D}_0(C^{-1}\otimes(B_{30}^{-1}\mathcal{H}_n B_{30}))], \\ \Delta v'\Psi_3\Delta\mathbf{y}_1 + \Delta v'\Pi_4 + \Delta v'\Phi_4\Delta v + \mathrm{tr}[\mathbf{W}_2\mathbf{D}_{-10}(C^{-1}\otimes(B_{30}^{-1}\mathcal{H}_n B_{30}))], \\ \Delta v'\Phi_5\Delta v - (T-1)\mathrm{tr}(G_{30}\mathcal{H}_n), \end{cases} \quad \text{(A.3)}$$

where all the quantities are defined in (3.3), except that $\mathcal{H}_n = \mathrm{diag}\{h_{ni}, i = \ldots, n\}$.

**Lemma A.8.** *Suppose Assumptions A-D hold for Model* (2.1). *If* $B_{10}^{-1}$ *and* $B_{30}^{-1}$ *are uniformly bounded in both row and column sums, we have* $\frac{1}{\sqrt{N}}S^{\circ}(\psi_0) \xrightarrow{D} N(0, \Gamma^{\circ})$.

**Proof of Lemma A.8.** First, under Assumptions C and D, the elements of all the $\Pi$ matrices are uniformly bounded.[8] By Lemma A.1, all the $\Phi$ and $\Psi$ matrices are uniformly bounded in both row and column sums. From (A.3), we see that $S^{\circ}(\psi_0)$ consists of three types of elements: $\Pi'\Delta v$, $\Delta v'\Phi\Delta v$ and $\Delta v'\Psi\Delta\mathbf{y}_1$, which can be written as

$$\Pi'\Delta v = \textstyle\sum_{t=1}^T \Pi_t^{*\prime} v_t, \ \ \Delta v\Phi\Delta v = \sum_{t=1}^T\sum_{s=1}^T v_t'\Phi_{ts}^* v_s, \ \text{ and } \Delta v'\Psi\Delta\mathbf{y}_1 = \sum_{t=1}^T v_t'\Psi_t^*\Delta y_1,$$

where $\Pi_t^*$, $\Phi_{ts}^*$ and $\Psi_t^*$ are formed by the elements of the partitioned $\Pi$, $\Phi$ and $\Psi$, respectively. By (2.1), $y_1 = B_{10}^{-1}B_{20}y_0 + \eta_1 + B_{10}^{-1}B_{30}^{-1}v_1$, leading to $\sum_{t=1}^T v_t'\Psi_t^*\Delta y_1 = \sum_{t=1}^T v_t'\Psi_t^{**}y_0 + \sum_{t=1}^T v_t'\Psi_t^{*+}v_1$, for suitably defined non-stochastic quantities $\eta_1$, $\Psi_t^{**}$ and $\Psi_t^{*+}$. These show that, for every non-zero $(p+5)\times 1$ vector of constants $c$, $c'S^{\circ}(\psi_0)$ can be expressed as

$$c'S^{\circ}(\psi_0) = \textstyle\sum_{t=1}^T\sum_{s=1}^T v_t'A_{ts}v_s + \sum_{t=1}^T v_t'B_t v_1 + \sum_{t=1}^T v_t'g(y_0) + c'\mu,$$

for suitably defined non-stochastic matrices $A_{ts}$ and $B_t$, the function $g(y_0)$ linear in $y_0$, and the non-stochastic vector $\mu$. As $\{y_0, v_1, \ldots, v_T\}$ are independent, the asymptotic normality of $\frac{1}{\sqrt{N}}c'S^{\circ}(\psi_0)$ follows from Lemma A.5. Finally, the Cramér-Wold devise leads to the joint asymptotic normality of $\frac{1}{\sqrt{N}}S^{\circ}(\psi_0)$. ■

**Recall:** for a square matrix $A$, $A^u$, $A^l$ and $A^d$ are, respectively, its upper-triangular, lower-triangular, and diagonal matrix such that $A = A^u + A^l + A^d$; $\Pi_t$, $\Phi_{ts}$ and $\Psi_{ts}$ the submatrices

[8]We omit the detailed discussions on the exact magnitude of the elements of $\Pi$ related to the degree of spatial dependence $\iota_n$ as specified in Assumption D.

of $\Pi$, $\Phi$ and $\Psi$ partitioned according to $t, s = 2, \ldots, T$; $\Psi_{t+} = \sum_{s=2}^{T} \Psi_{ts}$, $t = 2, \ldots, T$, $\Theta = \Psi_{2+}(B_{30}B_{10})^{-1}$, $\Delta y_1^\circ = B_{30}B_{10}\Delta y_1$, and $\Delta y_{1t}^* = \Psi_{t+}\Delta y_1$; $\mathcal{F}_{n,i} = \mathcal{F}_{n,0} \otimes \mathcal{G}_{n,i}$, with $\{\mathcal{G}_{n,i}\}$ being the increasing sequence of $\sigma$-fields generated by $(v_{j1}, \ldots, v_{jT}, j = 1, \ldots, i), i = 1, \ldots, n$, and $\mathcal{F}_{n,0}$ being the $\sigma$-field generated by $(v_0, \Delta y_0)$.

**Lemma A.9.** *Suppose Assumptions A-D hold for Model* (2.1)*. Consider the key quantities in* $S^\circ(\psi_0)$*:* $(\Pi'\Delta v,\ \Delta v'\Phi\Delta v,\ \Delta v'\Psi\Delta \mathbf{y}_1)' \equiv Q(\psi_0)$*. Define*

$$\begin{aligned} g_{1i} &= \textstyle\sum_{t=2}^{T} \Pi_{it}'\Delta v_{it}, \\ g_{2i} &= \textstyle\sum_{t=2}^{T} (\Delta v_{it}\Delta\xi_{it} + \Delta v_{it}\Delta v_{it}^* - \sigma_{v0}^2 d_{it}), \\ g_{3i} &= \Delta v_{2i}\Delta\zeta_i + \Theta_{ii}(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni}) + \textstyle\sum_{t=3}^{T} \Delta v_{it}\Delta y_{1it}^*, \end{aligned}$$

*where* $\xi_t = \sum_{s=2}^{T}(\Phi_{st}^{u\prime} + \Phi_{ts}^{l})\Delta v_s$*,* $\Delta v_t^* = \sum_{s=2}^{T}\Phi_{ts}^{d}\Delta v_s$*, and* $\{d_{it}\}$ *are the diagonal elements of* $\Phi(C \otimes \mathcal{H}_n)$*,* $\{\Delta\zeta_i\} = \Delta\zeta = (\Theta^u + \Theta^l)\Delta y_1^\circ$*, and* $\text{diag}\{\Theta_{ii}\} = \Theta^d$*. We have,*

(*i*) $Q(\psi_0) - \mathrm{E}[Q(\psi_0)] = \sum_{i=1}^{n} \mathbf{g}_i$*, where* $\mathbf{g}_i = (g_{1i}, g_{2i}, g_{3i})'$*,*
(*ii*) $\text{Var}[Q(\psi_0)] = \sum_{i=1}^{n} \mathrm{E}(\mathbf{g}_i\mathbf{g}_i')$*, and*
(*iii*) $\frac{1}{N}\sum_{i=1}^{n}[\mathbf{g}_i\mathbf{g}_i' - \mathrm{E}(\mathbf{g}_i\mathbf{g}_i')] = o_p(1)$*.*

**Proof of Lemma A.9.** Proof of (*i*) is trivial, and the result (*ii*) follows from the fact that $\{\mathbf{g}_i, \mathcal{F}_{n,i}\}$ forms a vector M.D. sequence. We focus on the proof of (*iii*). To facilitate the proof, the following *dot* notation is convenient: (*a*) for the $N \times 1$ vector $\Delta v$ with elements $\{\Delta v_{it}\}$ double indexed by $i = 1, \ldots, n$ for each $t = 2, \ldots, T$, $\Delta v_{\cdot t}$ is the subvector that contains all the elements with the same $t$, and $\Delta v_{i\cdot}$ is the subvector that picks up the elements with the same $i$; (*b*) for the $N \times N$ matrix $\Phi$ with elements $\{\Phi_{it,js},\ i,j = 1, \ldots, n;\ t,s = 2, \ldots, T\}$, where $it$ is the double index for the rows and $js$ the double index for the columns, $\Phi_{\cdot t,\cdot s}$ is the $n \times n$ submatrix corresponding to the $(t,s)$ periods, $\Phi_{i\cdot,j\cdot}$ the $(T-1) \times (T-1)$ submatrix corresponding to the $(i,j)$ units, $\Phi_{it,j\cdot}$ the $(T-1) \times 1$ subvector that picks up the element from the $it$th row corresponding to $s = 2, \ldots, T$.

With the vector dot notation, we have $g_{1i} = \Pi_{i\cdot}'\Delta v_{i\cdot}$, $g_{2i} = \Delta v_{i\cdot}'\Delta\xi_{i\cdot} + \Delta v_{i\cdot}'\Delta v_{i\cdot}^* - 1_{T-1}'d_{i\cdot}$, and $g_{3i} = \Delta v_{2i}\Delta\zeta_i + \Theta_{ii}(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni}) + \Delta v_{i-}'\Delta y_{1i-}^*$, where $d_{it}$ are the diagonal elements of $\Phi(C \times \mathcal{H})$ and '$_-$' plays the same role as '.' but corresponds to $t = 3, \ldots, T$. Note that under Assumptions D and E, one can easily see by Lemmas A.1-A.3 that the elements of all the $\Pi$'s, $\Phi$'s,, and $\Psi$'s are all uniformly bounded. The proofs proceed by applying the weak law of large numbers (WLLN) for M.D. arrays, see, e.g., Davidson (1994, p. 299).

First, with $g_{1i}^\circ = \Pi_{i\cdot}'\Delta v_{i\cdot}$, $\frac{1}{N}\sum_{i=1}^{n} g_{1i}g_{1i}' - \mathrm{E}(g_{1i}g_{1i}')] = \frac{1}{N}\sum_{i=1}^{n}\Pi_{i\cdot}'(\Delta v_{i\cdot}\Delta v_{i\cdot}' - \sigma_{v0}^2 C)\Pi_{i\cdot} \equiv \frac{1}{N}\sum_{i=1}^{n} U_{n,i}$, where $C$ is defined below (2.3). Without loss of generality, assume $U_{ni}$ is a scalar, as if not we can work on each element of it. Clearly, $\{U_{n,i}\}$ are independent, thus form a M.D. array. By Assumption B and using the fact that the elements of $\Pi_{i\cdot}$ are uniformly bounded, it is easy to show that $\mathrm{E}|U_{n,i}|^{1+\epsilon} \leq K_u < \infty$, for $\epsilon > 0$. Thus, $\{U_{n,i}\}$ are uniformly integrable. With the constant coefficients $\frac{1}{N}$ the other two conditions of WLLN for M.D. arrays of Davidson are satisfied. Thus, $\frac{1}{N}\sum_{i=1}^{n} U_{n,i} \xrightarrow{p} 0$.

Second, with $g_{2i} = \Delta v_{i\cdot}' \Delta \xi_{i\cdot} + \Delta v_{i\cdot}' \Delta v_{i\cdot}^* - 1_{T-1}' d_{i\cdot}$, we have,

$$\begin{aligned} & \tfrac{1}{N}\textstyle\sum_{i=1}^n [g_{2i}^2 - \mathrm{E}(g_{2i}^2)] \\ = \; & \tfrac{1}{N}\textstyle\sum_{i=1}^n [(\Delta v_{i\cdot}' \Delta \xi_{i\cdot})^2 - \mathrm{E}((\Delta v_{i\cdot}' \Delta \xi_{i\cdot})^2)] + \tfrac{1}{N}\sum_{i=1}^n [(\Delta v_{i\cdot}' \Delta v_{i\cdot}^*)^2 - \mathrm{E}((\Delta v_{i\cdot}' \Delta v_{i\cdot}^*)^2)] \\ & + \tfrac{2}{N}\textstyle\sum_{i=1}^n (\Delta v_{i\cdot}' \Delta \xi_{i\cdot})(\Delta v_{i\cdot}' \Delta v_{i\cdot}^*) - \tfrac{2}{N}\sum_{i=1}^n (1_{T-1}' d_{i\cdot})(\Delta v_{i\cdot}' \Delta \xi_{i\cdot}) \\ & - \tfrac{2}{N}\textstyle\sum_{i=1}^n [(1_{T-1}' d_{i\cdot})(\Delta v_{i\cdot}' \Delta v_{i\cdot}^* - \mathrm{E}(\Delta v_{i\cdot}' \Delta v_{i\cdot}^*))] \equiv \sum_{r=1}^5 H_r. \end{aligned}$$

Now, $H_1 = \frac{1}{N}\sum_{i=1}^n [\Delta \xi_{i\cdot}' (\Delta v_{i\cdot} \Delta v_{i\cdot}' - \sigma_{vo}^2 C)\Delta \xi_{i\cdot}] + \frac{\sigma_{v0}^2}{N}\sum_{i=1}^n [\Delta \xi_{i\cdot}' C \Delta \xi_{i\cdot} - \mathrm{E}(\Delta \xi_{i\cdot}' C \Delta \xi_{i\cdot})]$. For the first term, letting $V_{n,i} = \Delta \xi_{i\cdot}' (\Delta v_{i\cdot} \Delta v_{i\cdot}' - \sigma_{vo}^2 C)\Delta \xi_{i\cdot}$, we have $\mathrm{E}(V_{n,i}|\mathcal{G}_{n,i-1}) = 0$ due to the fact that $\Delta \xi_{i\cdot}$ is $\mathcal{G}_{n,i-1}$-measurable. Thus, $\{V_{n,i}, \mathcal{G}_{n,i}\}$ form a M.D. array. It is easy to see that $\mathrm{E}|V_{n,i}^{1+\epsilon}| \le K_v < \infty$, for some $\epsilon > 0$. Thus, $\{V_{n,i}\}$ is uniformly integrable. The other two conditions of the WLLN for M.D. arrays of Davidson are satisfied. Thus, $\frac{1}{N}\sum_{i=1}^n V_{n,i} \xrightarrow{p} 0$.

For the second term of $H_1$, recall $\xi_t = \sum_{s=2}^T (\Phi_{ts}^{w\prime} + \Phi_{ts}^{\ell})\Delta v_s$. We have,

$$\Delta \xi_{it} = \textstyle\sum_{s=2}^T \sum_{j=1}^{i-1} (\Phi_{jt,is} + \Phi_{it,js})\Delta v_{js} = \sum_{j=1}^{i-1}\sum_{s=2}^T (\Phi_{jt,is} + \Phi_{it,js})\Delta v_{js} = \sum_{j=1}^{i-1} \phi_{ijt} \Delta v_{j\cdot},$$

where $\phi_{ijt} = (\Phi_{ji,t\cdot} + \Phi_{ij,t\cdot})$. Thus, $(\Delta \xi_{it})^2 - \mathrm{E}[(\Delta \xi_{it})^2] = \sum_{j=1}^{i-1} [\phi_{ijt\cdot}' (\Delta v_{j\cdot} \Delta v_{j\cdot}' - \sigma_{v0}^2 C)\phi_{ijt\cdot}] + 2\sum_{j=1}^{i-1}\sum_{k=1}^{j-1} \Delta v_{j\cdot}' \phi_{ijt\cdot} \phi_{ikt\cdot}' \Delta v_{k\cdot}$. It follows that

$$\begin{aligned} & \tfrac{1}{N}\textstyle\sum_{i=1}^n \{(\Delta \xi_{it})^2 - \mathrm{E}[(\Delta \xi_{it})^2]\} \\ = \; & \tfrac{1}{N}\textstyle\sum_{i=1}^n \sum_{j=1}^{i-1} [\phi_{ijt\cdot}' (\Delta v_{j\cdot} \Delta v_{j\cdot}' - \sigma_{v0}^2 C)\phi_{ijt\cdot}] + 2\tfrac{1}{N}\sum_{i=1}^n \sum_{j=1}^{i-1}\sum_{k=1}^{j-1} \Delta v_{j\cdot}' \phi_{ijt\cdot} \phi_{ikt\cdot}' \Delta v_{k\cdot} \\ = \; & \tfrac{1}{N}\textstyle\sum_{j=1}^{n-1} \{ \sum_{i=j+1}^n [\phi_{ijt\cdot}' (\Delta v_{j\cdot} \Delta v_{j\cdot}' - \sigma_{v0}^2 C)\phi_{ijt\cdot}]\} \\ & + 2\tfrac{1}{N}\textstyle\sum_{j=1}^{n-1} \Delta v_{j\cdot}' \{ \sum_{i=j+1}^n \sum_{k=1}^{j-1} \phi_{ijt\cdot} \phi_{ikt\cdot}' \Delta v_{k\cdot} \}. \end{aligned}$$

Clearly, the first term is the 'average' of $n-1$ independent terms, and the second is the 'average' of a M.D. array as the term in the curling brackets is $G_{n,j-1}$-measurable. Conditions of Theorem 19.7 of Davidson (1994) are easily verified, and hence $\frac{1}{N}\sum_{i=1}^n \{(\Delta \xi_{it})^2 - \mathrm{E}[(\Delta \xi_{it})^2]\} = o_p(1)$. Similarly, one shows that $\frac{1}{N}\sum_{i=1}^n \{\Delta \xi_{it} \Delta \xi_{is} - \mathrm{E}[(\Delta \xi_{it} \Delta \xi_{is})]\} = o_p(1)$ for $s \ne t$. Thus, $\frac{\sigma_{v0}^2}{N}\sum_{i=1}^n [\Delta \xi_{i\cdot}' C \Delta \xi_{i\cdot} - \mathrm{E}(\Delta \xi_{i\cdot}' C \Delta \xi_{i\cdot})] = o_p(1)$, and $H_1 = o_p(1)$.

The proofs for $H_3$ and $H_4$ can be done in a similar manner as the proof for the second term of $H_1$. The proofs for $H_2$ and $H_5$ are similar to the proof of the first part of $H_1$, as they each involves a sum of $n$ independent terms.

Third, with $g_{3i} = \Delta v_{2i} \Delta \zeta_i + \Theta_{ii}(\Delta v_{2i} \Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni}) + \Delta v_{i-}' \Delta y_{1i-}^*$, we obtain,

$$\begin{aligned} & \tfrac{1}{N}\textstyle\sum_{i=1}^n [g_{3i}^2 - \mathrm{E}(g_{3i}^2)] = \tfrac{1}{N}\sum_{i=1}^n [(\Delta v_{2i}^2 - 2\sigma_{v0}^2 h_{ni})\Delta \zeta_i^2] + \tfrac{2\sigma_{v0}^2}{N}\sum_{i=1}^n h_{ni}[\Delta \zeta_i^2 - \mathrm{E}(\Delta \zeta_i^2)] \\ & \quad + \tfrac{1}{N}\textstyle\sum_{i=1}^n \Theta_{ii}^2 [(\Delta v_{2i} \Delta y_{1i}^\circ)^2 - \mathrm{E}((\Delta v_{2i} \Delta y_{1i}^\circ)^2)] \\ & \quad + \tfrac{2\sigma_{v0}^2}{N}\textstyle\sum_{i=1}^n \Theta_{ii}^2 h_{ni}[\Delta v_{2i} \Delta y_{1i}^\circ - \mathrm{E}(\Delta v_{2i} \Delta y_{1i}^\circ)] \\ & \quad + \tfrac{1}{N}\textstyle\sum_{i=1}^n [(\Delta v_{i-}' \Delta y_{1i-}^*)^2 - \mathrm{E}((\Delta v_{i-}' \Delta y_{1i-}^*)^2)] \\ & \quad + \tfrac{2}{N}\textstyle\sum_{i=1}^n \Theta_{ii}[\Delta v_{2i}^2 \Delta \zeta_i \Delta y_{1i}^\circ - \mathrm{E}(\Delta v_{2i}^2 \Delta \zeta_i \Delta y_{1i}^\circ)] + \tfrac{2\sigma_{v0}^2}{N}\sum_{i=1}^n \Theta_{ii} h_{ni} \Delta v_{2i} \Delta \zeta_i \\ & \quad + \tfrac{2}{N}\textstyle\sum_{i=1}^n [\Delta v_{2i} \Delta \zeta_i (\Delta v_{i-}' \Delta y_{1i-}^*) - \mathrm{E}(\Delta v_{2i} \Delta \zeta_i (\Delta v_{i-}' \Delta y_{1i-}^*))] \\ & \quad + \tfrac{2}{N}\textstyle\sum_{i=1}^n \Theta_{ii}[(\Delta v_{2i} \Delta y_{1i}^\circ)(\Delta v_{i-}' \Delta y_{1i-}^*) - \mathrm{E}((\Delta v_{2i} \Delta y_{1i}^\circ)(\Delta v_{i-}' \Delta y_{1i-}^*))] \\ & \quad + \tfrac{2\sigma_{v0}^2}{N}\textstyle\sum_{i=1}^n \Theta_{ii} h_{ni}[\Delta v_{i-}' \Delta y_{1i-}^* - \mathrm{E}(\Delta v_{i-}' \Delta y_{1i-}^*)] \equiv \sum_{r=1}^{10} Q_r. \end{aligned}$$

As $\Delta \zeta_i^2$ is $\mathcal{F}_{n,i-1}$-measurable, $Q_1$ is the average of a M.D. array and its convergence to 0 in probability follows from WLLN for M.D. array. For $Q_2$, note that $\Delta\zeta = (\Theta^u + \Theta^\ell)\Delta y_1^\circ = (\Theta^u + \Theta^\ell)B_{30}B_{10}\Delta y_1$. It follows that $Q_2 = \frac{2\sigma_{v0}^2}{N}\sum_{i=1}^n(\Delta y_1' A \Delta y_1 - \mathrm{E}(\Delta y_1' A \Delta y_1)] = o_p(1)$ by Assumption F, where $A = ((\Theta^u + \Theta^\ell)B_{30}B_{10})'(\Theta^{u\prime} + \Theta^\ell)B_{30}B_{10}$ is easily seen to be uniformly bounded in both row and column sums. Writing $\Delta y_1^\circ = B_{30}B_{10}\Delta y_0 + B_{30}\Delta x_1\beta_0 + \Delta v_1 \equiv g(y_0, v_0) + v_1$, the convergence of $Q_3$, $Q_4$ and $Q_6$ can be easily proved though tedious. The results for $Q_5$ and $Q_{10}$ are proved by the independence between $\Delta v_{i-}'$ and $\Delta y_{1i-}^*$ are independent, $\Delta y_{1t}^* = \Phi_{t+}\Delta y_1$, and Assumption F. The convergence of $Q_7$ to 0 in probability follows that of $Q_1$. Finally, the results for $Q_8$ and $Q_9$ can be proved by further writing $\Delta y_{1t}^* = \Phi_{t+}\Delta y_1 = \Phi_{t+}(B_{30}B_{10})^{-1}\Delta y_1^\circ \equiv q(\Delta y_0, v_0) + \Phi_{t+}(B_{30}B_{10})^{-1}v_1$.

Subsequently, for the cross-product terms, we have,

$$
\begin{aligned}
& \tfrac{1}{N}\textstyle\sum_{i=1}^n[g_{1i}g_{2i} - \mathrm{E}(g_{1i}g_{2i})] \\
= \; & \tfrac{1}{N}\textstyle\sum_{i=1}^n[\Pi_{i\cdot}'(\Delta v_{i\cdot}\Delta v_{i\cdot}' - \sigma_{v0}^2 C)\Delta\xi_{i\cdot}] + \tfrac{\sigma_{v0}^2}{N}\textstyle\sum_{i=1}^n(\Pi_{i\cdot}' C\Delta\xi_{i\cdot}) \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n \Pi_{i\cdot}'[\Delta v_{i\cdot}\Delta v_{i\cdot}'\Delta v_{i\cdot}^* - \mathrm{E}(\Delta v_{i\cdot}\Delta v_{i\cdot}'\Delta v_{i\cdot}^*)] + \tfrac{1}{N}\textstyle\sum_{i=1}^n[(1_{T-1}' d_{i\cdot})\Pi_{i\cdot}'\Delta v_{i\cdot}].
\end{aligned}
$$

$$
\begin{aligned}
& \tfrac{1}{N}\textstyle\sum_{i=1}^n[g_{1i}g_{3i} - \mathrm{E}(g_{1i}g_{3i})] \\
= \; & \tfrac{1}{N}\textstyle\sum_{i=1}^n \Pi_{i\cdot}'[\Delta v_{i\cdot}\Delta v_{2i}\Delta\zeta_i - \mathrm{E}(\Delta v_{i\cdot}\Delta v_{2i}\Delta\zeta_i)] \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n \Theta_{ii}\Pi_{i\cdot}'[\Delta v_{i\cdot}(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni}) - \mathrm{E}(\Delta v_{i\cdot}(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni}))] \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n \Pi_{i\cdot}'[\Delta v_{i\cdot}\Delta v_{i-}'\Delta y_{1i-}^* - \mathrm{E}(\Delta v_{i\cdot}\Delta v_{i-}'\Delta y_{1i-}^*)].
\end{aligned}
$$

$$
\begin{aligned}
& \tfrac{1}{N}\textstyle\sum_{i=1}^n[g_{2i}g_{3i} - \mathrm{E}(g_{2i}g_{3i})] \\
= \; & \tfrac{1}{N}\textstyle\sum_{i=1}^n[(\Delta v_{i\cdot}'\Delta\xi_i)(\Delta v_{2i}\Delta\zeta_i) - \mathrm{E}((\Delta v_{i\cdot}'\Delta\xi_i)(\Delta v_{2i}\Delta\zeta_i))] \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n \Theta_{ii}[(\Delta v_{i\cdot}'\Delta\xi_i)(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni}) - \mathrm{E}((\Delta v_{i\cdot}'\Delta\xi_i)(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni}))] \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n[(\Delta v_{i\cdot}'\Delta\xi_i)(\Delta v_{i-}'\Delta y_{1i-}^*) - \mathrm{E}((\Delta v_{i\cdot}'\Delta\xi_i)(\Delta v_{i-}'\Delta y_{1i-}^*))] \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n[(\Delta v_{i\cdot}'\Delta v_{i\cdot}^*)(\Delta v_{2i}\Delta\zeta_i) - \mathrm{E}((\Delta v_{i\cdot}'\Delta v_{i\cdot}^*)(\Delta v_{2i}\Delta\zeta_i))] \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n[(\Delta v_{i\cdot}'\Delta v_{i\cdot}^*)(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni}) - \mathrm{E}((\Delta v_{i\cdot}'\Delta v_{i\cdot}^*)(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni}))] \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n[(\Delta v_{i\cdot}'\Delta v_{i\cdot}^*)(\Delta v_{i-}'\Delta y_{1i-}^*) - \mathrm{E}((\Delta v_{i\cdot}'\Delta v_{i\cdot}^*)(\Delta v_{i-}'\Delta y_{1i-}^*))] \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n[(1_{T-1}' d_{i\cdot})\Delta v_{2i}\Delta\zeta_i] + \tfrac{1}{N}\textstyle\sum_{i=1}^n[(1_{T-1}' d_{i\cdot})\Theta_{ii}(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2 h_{ni})] \\
& + \tfrac{1}{N}\textstyle\sum_{i=1}^n[(1_{T-1}' d_{i\cdot})(\Delta v_{i-}'\Delta y_{1i-}^* - \mathrm{E}(\Delta v_{i-}'\Delta y_{1i-}^*))]
\end{aligned}
$$

The convergence of each of the terms above can be proved in a similarly manner as these terms appear in similar forms as the terms appeared in the $H_r$ and $Q_r$. ■

## Appendix B: Proofs of the Main Theoretical Results

**Proof of Theorem 3.1.** The quantities needed for evaluating the AQS function defined in (3.3) become: $\Pi_1 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\Delta X$, $\Pi_2 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbb{B}_{-1}\Delta X\beta$, $\Pi_3 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_1\Delta X\beta$, $\Pi_4 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_2\mathbb{B}_{-1}\Delta X\beta$, $\Phi_1 = \frac{1}{2\sigma_{v0}^4}\mathbf{C}^{-1}$, $\Phi_2=\frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbb{B}_{-1}$, $\Phi_3 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_1$, $\Phi_4 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_2\mathbb{B}_{-1}$, $\Phi_5 = \frac{1}{2\sigma_{v0}^2}[C^{-1} \otimes (W_3' + W_3)]$, $\Psi_1 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbb{R}_{-1}$, $\Psi_2 = 0$, $\Psi_3 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_2\mathbb{R}_{-1}$, $\mathbb{R}_{-1} = \text{blkdiag}(I_n, 0, \ldots, 0)$, $\mathbb{B}_{-1} = I_{T-1}^* \otimes I_n$, and $I_{T-1}^*$ is a $(T-1)\times(T-1)$ matrix with elements 1 on the positions immediately below the diagonal elements, and zero elsewhere. Further, $\mathcal{B}_0 = 0_n$, and hence $\mathbf{D}_0 = -C \otimes I_n$ and $\mathbf{D}_{-10} = -C_{-1} \otimes I_n$, where

$$C_{-1} = \begin{pmatrix} -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 2 & -1 & 0 & \cdots & 0 & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 2 & -1 \end{pmatrix}_{(T-1)\times(T-1)}.$$

These show that with $\psi_0 = (\beta_0', \sigma_{v0}, 0, 0, 0, 0)'$, all the $\Phi$ and $\Psi$ matrices are either of the form $A \otimes I_n$ or of the form $A \times W$ for some $(T-1)\times(T-1)$ matrix $A$ and a spatial weight matrix $W$ satisfying Assumption D. Thus, $\text{E}[S^*(\psi_0)] = 0$ even when the errors are heteroskedastic by Lemma A.7. Hence by Lemma A.8, we have $\frac{1}{\sqrt{N}}S^*(\psi_0) \xrightarrow{D} N(0, \Gamma^*(\psi_0)]$.

By the mean value theorem, one easily shows that $\frac{1}{\sqrt{N}}[S_\delta^*(\tilde{\psi}) - S_\delta^*(\psi_0)] = o_p(1)$, where $\tilde{\psi} = (\tilde{\beta}', \tilde{\sigma}_{v0}^2, 0, 0, 0, 0)'$ and we note that the OLS estimators $\tilde{\beta}$ and $\tilde{\sigma}_{v0}^2$ are robust against unknown heteroskedasticity $\{h_{ni}\}$. Now, since by (3.11) $S^*(\psi_0) = \sum_{i=1}^n \mathbf{g}_i$, where $\{\mathbf{g}_i, \mathcal{F}_{n,i}\}$ form a vector M.D. sequence, we have $\frac{1}{N}\sum_{i=1}^n[\mathbf{g}_i\mathbf{g}_i' - \text{E}(\mathbf{g}_i\mathbf{g}_i')] = o_p(1)$ by Lemma A.9. Finally, by the mean value theorem and the consistency of $\tilde{\beta}$ and $\tilde{\sigma}_{v0}^2$, one shows that $\frac{1}{N}\sum_{i=1}^n(\tilde{\mathbf{g}}_i\tilde{\mathbf{g}}_i' - \mathbf{g}_i\mathbf{g}_i') = o_p(1)$ under heteroskedasticity. This completes the proof of Theorem 3.1. ∎

**Proof of Lemma 3.1.** Consider the AQS vector $S^*(\beta, \sigma_v^2, \rho)$ for the DPD model, and the concentrated AQS function which defines $\tilde{\rho}$ under $H_0^{\texttt{DPD}}$:

$$S_{\texttt{DPD}}^{*c}(\rho) = \tfrac{1}{\tilde{\sigma}_v^2(\rho)}\Delta\tilde{v}'(\rho)\mathbf{C}^{-1}\Delta Y_{-1} + n\big(\tfrac{1}{1-\rho} - \tfrac{1-\rho^T}{T(1-\rho)^2}\big),$$

where $\tilde{\beta}(\rho) = (\Delta X'\mathbf{C}^{-1}\Delta X)^{-1}\Delta X'\mathbf{C}^{-1}(\Delta Y - \rho\Delta Y_{-1})$ and $\tilde{\sigma}_v^2(\rho) = \frac{1}{N}\Delta\tilde{v}'(\rho)\mathbf{C}^{-1}\Delta\tilde{v}(\rho)$, where $\Delta\tilde{v}(\rho) = \Delta Y - \rho\Delta Y_{-1} - \Delta X\tilde{\beta}(\rho)$.

Define $\bar{S}^*(\beta, \sigma_v^2, \rho) = \text{E}[S^*(\beta, \sigma_v^2, \rho)]$. Given $\rho$, $\bar{S}^*(\beta, \sigma_v^2, \rho) = 0$ is partially solved at $\bar{\beta}(\rho) = (\Delta X'\mathbf{C}^{-1}\Delta X)^{-1}\Delta X'\mathbf{C}^{-1}(\text{E}\Delta Y - \rho\text{E}\Delta Y_{-1})$ and $\bar{\sigma}_v^2(\rho) = \frac{1}{N}\text{E}[\Delta\bar{v}(\rho)'\mathbf{C}^{-1}\Delta\bar{v}(\rho)]$, where $\Delta\bar{v}(\rho) = \Delta Y - \rho\Delta Y_{-1} - \Delta X\bar{\beta}(\rho)$. Substituting $\bar{\beta}(\rho)$ and $\bar{\sigma}_v^2(\rho)$ back into $\bar{S}^*(\beta, \sigma_v^2, \rho)$ gives the population counter part of $S_{\texttt{DPD}}^{*c}(\rho)$ as

$$\bar{S}_{\texttt{DPD}}^{*c}(\rho) = \tfrac{1}{\bar{\sigma}_v^2(\rho)}\text{E}[\Delta\bar{v}'(\rho)\mathbf{C}^{-1}\Delta Y_{-1}] + n\big(\tfrac{1}{1-\rho} - \tfrac{1-\rho^T}{T(1-\rho)^2}\big).$$

By Theorem 5.9 of van der Vaart (1998), $\tilde{\rho}$ will be consistent if $(i)$ $\inf_{\rho:|\rho-\rho_0|\geq\epsilon} |\bar{S}^{*c}_{\mathtt{DPD}}(\rho)| > 0$ for every $\epsilon > 0$, and $(ii)$ $\sup_{\rho\in\Upsilon} \frac{1}{\sqrt{N}}|S^{*c}_{\mathtt{DPD}}(\rho) - \bar{S}^{*c}_{\mathtt{DPD}}(\rho)| \xrightarrow{p} 0$, which are straightforward. The asymptotic normality can be proved using Lemma A.5. Details of the proof are available from the author upon request. ∎

**Proof of Theorem 3.2.** First, with $\psi_0 = (\beta_0, \sigma^2_{v0}, \rho_0, 0'_3)'$ it is easy to show that $\mathrm{E}[S^*(\psi_0)] = 0$ under the general heteroskedasticity $\{h_{ni}\}$. Thus, under $\lambda = 0$, $S^{\circ}(\psi_0)$ defined in (A.3) reduces to $S^*(\psi_0)$ defined in (3.3), and hence by Lemma A.8, one shows that $\frac{1}{\sqrt{N}}S^*(\psi_0) \xrightarrow{D} N(0, \Gamma^*(\psi_0))$. By Lemma A.9, one shows that $\frac{1}{N}\sum_{i=1}^n[\mathbf{g}_{n,i}\mathbf{g}'_{n,i} - \mathrm{E}(\mathbf{g}_i\mathbf{g}'_i)] \xrightarrow{p} 0$. By the mean value theorem, and $\sqrt{N}$ consistency and robustness of $\tilde{\beta}$, $\tilde{\sigma}^2_v$ and $\tilde{\rho}$ against unknown heteroskedasticity $\{h_{ni}\}$ as shown in Lemma 3.1, we have $\frac{1}{\sqrt{N}}[S^*_\lambda(\tilde{\psi}) - S^*_\lambda(\psi_0)] \xrightarrow{p} 0$ where $\tilde{\psi} = (\tilde{\beta}', \tilde{\sigma}^2_v, \tilde{\rho}, 0_3)'$, and $\frac{1}{N}\sum_{i=1}^n(\tilde{\mathbf{g}}_{n,i}\tilde{\mathbf{g}}'_{n,i} - \mathbf{g}_i\mathbf{g}'_i) \xrightarrow{p} 0$. The result of Theorem 3.2 thus follows. ∎

**Proof of Theorem 3.3.** Referring to the AQS vector $S^*(\psi_0)$ given in (3.3) setting $\rho_0$ to 0, Lemma A.1 shows that all the $\Phi$ and $\Psi$ matrices are uniformly bounded in both row and column sums. Assumptions C and D, and the additional assumptions stated in the theorem guarantee that the elements of all the $\Pi$ quantities are uniformly bounded. Lemma A.5 and hence Lemma A.8 are applicable under homoskedastic errors. Now, $S^{\circ}(\psi_0)$ defined in (A.3) reduces to $S^*(\psi_0)$ defined in (3.3), and hence Lemma A.8 leads to $\frac{1}{\sqrt{N}}S^*(\psi_0) \xrightarrow{D} N(0, \Gamma^*(\psi_0))$. With the $\sqrt{N}$-consistency of $\tilde{\psi} = (\tilde{\beta}', \tilde{\sigma}^2_v, 0, \tilde{\lambda})'$, we have $\frac{1}{\sqrt{N}}[S^*_\rho(\tilde{\psi}) - S^*_\rho(\psi_0)] \xrightarrow{p} 0$ by the mean value theorem and Lemma A.2. Now, by Lemma A.9 with $h_{ni} = 1, i = 1, \ldots, n$, we have $\frac{1}{N}\sum_{i=1}^n[\mathbf{g}_{n,i}\mathbf{g}'_{n,i} - \mathrm{E}(\mathbf{g}_i\mathbf{g}'_i)] \xrightarrow{p} 0$. By the mean value theorem, $\sqrt{N}$-consistency of $\tilde{\psi}$, and Lemma A.2, we have $\frac{1}{N}\sum_{i=1}^n(\tilde{\mathbf{g}}_{n,i}\tilde{\mathbf{g}}'_{n,i} - \mathbf{g}_i\mathbf{g}'_i) \xrightarrow{p} 0$. The result of the theorem thus follows. ∎

**Proof of Theorems 3.4-3.7.** Similar to the proof of Theorem 3.3. The details are available from the author upon request. ∎

**Proof of Corollaries 4.1-4.7.** Lemma A.3, Assumptions C and D, and the additional assumptions (if any) stated in the corollaries guarantee that the matrices $\mathbf{M}^*_0$ and $\mathbf{M}^{**}_0$ appeared in (4.2) and (4.4) are both uniformly bounded in both row and column sums. The rest of the proofs of these corollaries parallel the proofs of Theorems 3.1-3.7. The details are available from the author upon request. ∎

A final note for the proofs of all the theorems and corollaries, the exact order of the key quantities in connection with the degree of spatial dependence represented by $\iota_n$ as in Assumption D can be learned through Lemma A.4 and Lemma A.5.

## References


[1] Anderson, T. W., Hsiao, C., 1981. Estimation of dynamic models with error components. *Journal of American Statistical Association* **76**, 598-606.

[2] Anderson, T. W., Hsiao, C., 1982. Formulation and estimation of dynamic models using panel data. *Journal of Econometrics* **18**, 47-82.

[3] Anselin, L., 2001. Rao's score test in spatial econometrics. *Journal of Statistical Planning and Inference* **97**, 113-139.

[4] Anselin L., Bera, A. K., 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: *Handbook of Applied Economic Statistics, Edited by Aman Ullah and David E. A. Giles.* New York: Marcel Dekker.

[5] Baltagi, B. H., Song, S. H., Koh W., 2003. Testing panel data regression models with spatial error correlation. *Journal of Econometrics* **117**, 123-150.

[6] Baltagi, B. H., Song, S. H., Jung, Koh W., 2007. Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics* **140**, 5-51.

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

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

[9] Bhargava, A., Sargan, J. D., 1983. Estimating dynamic random effects models from panel data covering short time periods. *Econometrica* 51, 1635-1659.

[10] Binder, M., Hsiao, C., Pesaran, M. H., 2005. Estimation and inference in short panel vector autoregressions with unit roots and cointegration. *Econometric Theory* 21, 795-837.

[11] Born, B., Breitung, J., 2011. Simple regression based tests for spatial dependence. *Econometrics Journal*, **14**, 330-342.

[12] Bun, M.J., Carree, M.A., 2005. Bias-corrected estimation in dynamic panel data models. *Journal of Business and Economic Statistics* 23, 200-210.

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

[14] Debarsy, N., Ertur, C., 2010. Testing for spatial autocorrelation in a fixed effects panel data model. *Regional Science and Urban Economics*, **40**, 453–70.

[15] Gouriéroux, C. Phillips, P. C. B, Yu, J., 2010. Indirect inference for dynamic panel models. *Journal of Econometrics* 157, 68-77.

[16] Hahn, J., Kuersteiner, G., 2002. Asymptotically unbiased inference for a dynamic model with fixed effects when both $n$ and $T$ are large. *Econometrica* **70**, 1639-1657.

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

[18] Hsiao, C., 2003. *Analysis of Panel Data. 2nd edition.* Cambridge University Press.

[19] Jin, F., Lee, L. F., 2015. On the bootstrap for Moran's I test for spatial dependence. *Journal of Econometrics* **184**, 295-314.

[20] 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.

[21] 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.

[22] Kruiniger, H., 2013. Quasi ML estimation of the panel AR(1) model with arbitrary initial conditions. *Journal of Econometrics* 173, 175-188.

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

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

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

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

[27] Robinson, P.M., Rossi, F., 2014. Improved Lagrange multiplier tests in spatial autoregressions. *Econometrics Journal* **17**, 139-164.

[28] Robinson, P.M., Rossi, F., 2015a. Refined tests for spatial correlation. *Econometric Theory* **31**, 1249-1280.

[29] Robinson, P.M., Rossi, F., 2015b. Refinements in maximum likelihood inference on spatial autocorrelation in panel data. *Journal of Econometrics* **189**, 447-456.

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

[31] van der Vaart, A. W., 1998. *Asymptotic Statistics.* Cambridge University Press.

[32] Yang, Z. L., 2010. A robust LM test for spatial error components. *Regional Science and Urban Economics* **40**, 299-310.

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

[34] Yang, Z. L., 2016. Unified M-estimation of fixed-effects spatial dynamic models with short panels. *Working Paper*, Singapore Management University.

**Table 1** Empirical Mean, sd and Size of $T^{\texttt{PD}}_{\texttt{AQS}}$ for Testing $\delta = 0$

| | | `AQS Test` | | | | | `SAQS Test` | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | dgp | mean | sd | 10% | 5% | 1% | mean | sd | 10% | 5% | 1% |
| | | `Group Interaction` | | | | | | | | | |
| 50 | 1 | 4.80 | 3.20 | .1590 | .0908 | .0206 | 4.15 | 2.62 | .0990 | .0378 | .0054 |
| | 2 | 6.01 | 4.11 | .2712 | .1742 | .0638 | 4.13 | 2.45 | .0880 | .0330 | .0024 |
| | 3 | 5.35 | 3.56 | .2156 | .1270 | .0352 | 4.14 | 2.49 | .0874 | .0356 | .0024 |
| 100 | 1 | 4.55 | 3.19 | .1448 | .0816 | .0192 | 4.17 | 2.79 | .1096 | .0552 | .0098 |
| | 2 | 5.24 | 3.67 | .2084 | .1240 | .0364 | 4.07 | 2.56 | .0928 | .0386 | .0040 |
| | 3 | 4.90 | 3.45 | .1792 | .1018 | .0288 | 4.17 | 2.72 | .1072 | .0494 | .0078 |
| 200 | 1 | 4.29 | 2.95 | .1258 | .0606 | .0142 | 4.10 | 2.76 | .1058 | .0478 | .0082 |
| | 2 | 4.79 | 3.43 | .1646 | .1008 | .0264 | 4.11 | 2.71 | .1002 | .0484 | .0070 |
| | 3 | 4.46 | 3.21 | .1396 | .0756 | .0214 | 4.06 | 2.76 | .1014 | .0490 | .0082 |
| 500 | 1 | 4.15 | 2.99 | .1128 | .0624 | .0136 | 4.07 | 2.90 | .1038 | .0572 | .0104 |
| | 2 | 4.33 | 3.13 | .1282 | .0672 | .0182 | 4.04 | 2.80 | .1018 | .0498 | .0076 |
| | 3 | 4.23 | 3.01 | .1218 | .0656 | .0150 | 4.05 | 2.82 | .1030 | .0522 | .0102 |
| | | `Rook Contiguity` | | | | | | | | | |
| 50 | 1 | 4.63 | 3.14 | .1496 | .0828 | .0180 | 3.99 | 2.56 | .0870 | .0356 | .0054 |
| | 2 | 5.79 | 3.94 | .2498 | .1568 | .0544 | 3.97 | 2.35 | .0720 | .0262 | .0016 |
| | 3 | 5.26 | 3.64 | .2046 | .1202 | .0352 | 4.04 | 2.50 | .0830 | .0340 | .0042 |
| 100 | 1 | 4.38 | 3.05 | .1290 | .0696 | .0150 | 4.02 | 2.70 | .0954 | .0458 | .0066 |
| | 2 | 5.06 | 3.65 | .1886 | .1106 | .0354 | 3.94 | 2.53 | .0818 | .0374 | .0046 |
| | 3 | 4.60 | 3.29 | .1540 | .0870 | .0234 | 3.92 | 2.60 | .0872 | .0374 | .0044 |
| 200 | 1 | 4.22 | 2.96 | .1190 | .0638 | .0132 | 4.03 | 2.76 | .1012 | .0510 | .0078 |
| | 2 | 4.75 | 3.42 | .1634 | .0926 | .0272 | 4.09 | 2.73 | .1004 | .0474 | .0082 |
| | 3 | 4.30 | 3.05 | .1216 | .0622 | .0164 | 3.93 | 2.66 | .0888 | .0388 | .0066 |
| 500 | 1 | 4.05 | 2.88 | .1074 | .0560 | .0104 | 3.98 | 2.79 | .1004 | .0506 | .0090 |
| | 2 | 4.28 | 3.01 | .1236 | .0626 | .0134 | 4.00 | 2.71 | .0950 | .0446 | .0070 |
| | 3 | 4.13 | 2.95 | .1152 | .0600 | .0118 | 3.96 | 2.77 | .1006 | .0472 | .0064 |
| | | `Group Interaction, Heteroskedastic Errors` | | | | | | | | | |
| 50 | 1 | 5.60 | 3.71 | .2348 | .1376 | .0430 | 4.22 | 2.48 | .0922 | .0336 | .0034 |
| | 2 | 7.16 | 4.93 | .3618 | .2608 | .1136 | 4.18 | 2.35 | .0770 | .0284 | .0026 |
| | 3 | 6.31 | 4.30 | .3030 | .2048 | .0764 | 4.19 | 2.45 | .0886 | .0316 | .0018 |
| 100 | 1 | 4.83 | 3.38 | .1652 | .0954 | .0278 | 4.15 | 2.69 | .1006 | .0464 | .0060 |
| | 2 | 5.73 | 4.18 | .2446 | .1580 | .0542 | 4.08 | 2.49 | .0886 | .0356 | .0016 |
| | 3 | 5.17 | 3.65 | .2030 | .1196 | .0336 | 4.10 | 2.57 | .0910 | .0376 | .0048 |
| 200 | 1 | 4.45 | 3.11 | .1404 | .0734 | .0164 | 4.10 | 2.75 | .1030 | .0462 | .0074 |
| | 2 | 5.09 | 3.74 | .1920 | .1156 | .0372 | 4.07 | 2.65 | .0926 | .0418 | .0054 |
| | 3 | 4.68 | 3.35 | .1582 | .0834 | .0216 | 4.06 | 2.69 | .0908 | .0428 | .0082 |
| 500 | 1 | 4.24 | 3.03 | .1180 | .0612 | .0154 | 4.06 | 2.83 | .1024 | .0512 | .0118 |
| | 2 | 4.62 | 3.37 | .1592 | .0848 | .0220 | 4.09 | 2.80 | .1068 | .0488 | .0082 |
| | 3 | 4.41 | 3.17 | .1352 | .0734 | .0190 | 4.07 | 2.80 | .1044 | .0490 | .0080 |

**Note**: for dgp, 1=normal, 2=normal mixture, and 3=lognormal.
for the last panel, heteroskedasticity $h_{ni} \propto |\frac{1}{T-1}\sum_{t=2}^{T}\Delta X_{it}|$.

**Table 2a** Empirical Mean, sd and Size of $T_{\texttt{AQS}}^{\texttt{DPD}}$ for Testing $\lambda = 0$, Homoskedastic Errors.

| | | AQS Test | | | | | SAQS Test | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | dgp | mean | sd | 10% | 5% | 1% | mean | sd | 10% | 5% | 1% |
| | | Group Interaction, $\rho = 0$ | | | | | | | | | |
| 50 | 1 | 3.72 | 2.90 | .1624 | .0902 | .0236 | 3.26 | 2.42 | .1116 | .0554 | .0092 |
| | 2 | 4.60 | 3.64 | .2590 | .1576 | .0566 | 3.35 | 2.36 | .1158 | .0520 | .0086 |
| | 3 | 4.10 | 3.11 | .2064 | .1214 | .0356 | 3.29 | 2.32 | .1094 | .0470 | .0060 |
| 100 | 1 | 3.45 | 2.79 | .1398 | .0774 | .0184 | 3.19 | 2.49 | .1118 | .0548 | .0116 |
| | 2 | 3.94 | 3.27 | .1944 | .1148 | .0374 | 3.20 | 2.41 | .1072 | .0526 | .0076 |
| | 3 | 3.76 | 3.12 | .1700 | .1012 | .0306 | 3.27 | 2.53 | .1200 | .0590 | .0122 |
| 200 | 1 | 3.20 | 2.65 | .1222 | .0642 | .0128 | 3.07 | 2.49 | .1088 | .0546 | .0090 |
| | 2 | 3.59 | 2.88 | .1556 | .0862 | .0218 | 3.17 | 2.40 | .1052 | .0520 | .0084 |
| | 3 | 3.38 | 2.72 | .1372 | .0704 | .0178 | 3.13 | 2.42 | .1060 | .0506 | .0102 |
| 500 | 1 | 3.04 | 2.50 | .1018 | .0530 | .0108 | 2.99 | 2.43 | .0940 | .0474 | .0100 |
| | 2 | 3.20 | 2.67 | .1188 | .0632 | .0136 | 3.02 | 2.42 | .0982 | .0506 | .0082 |
| | 3 | 3.15 | 2.54 | .1132 | .0606 | .0104 | 3.04 | 2.41 | .1034 | .0496 | .0074 |
| | | Rook Contiguity, $\rho = 0$ | | | | | | | | | |
| 50 | 1 | 3.56 | 2.84 | .1504 | .0802 | .0230 | 3.11 | 2.36 | .1022 | .0462 | .0068 |
| | 2 | 4.45 | 3.53 | .2406 | .1556 | .0498 | 3.22 | 2.28 | .1034 | .0444 | .0048 |
| | 3 | 3.98 | 3.16 | .2000 | .1178 | .0328 | 3.16 | 2.29 | .1048 | .0442 | .0046 |
| 100 | 1 | 3.26 | 2.67 | .1212 | .0644 | .0156 | 3.03 | 2.38 | .0980 | .0486 | .0076 |
| | 2 | 3.77 | 3.04 | .1708 | .0988 | .0272 | 3.07 | 2.27 | .0910 | .0416 | .0064 |
| | 3 | 3.49 | 2.83 | .1460 | .0810 | .0214 | 3.05 | 2.33 | .0996 | .0420 | .0074 |
| 200 | 1 | 3.09 | 2.55 | .1050 | .0534 | .0148 | 2.97 | 2.41 | .0946 | .0464 | .0096 |
| | 2 | 3.40 | 2.78 | .1362 | .0828 | .0172 | 3.01 | 2.34 | .0986 | .0460 | .0068 |
| | 3 | 3.28 | 2.75 | .1252 | .0714 | .0186 | 3.04 | 2.45 | .1032 | .0508 | .0094 |
| 500 | 1 | 3.01 | 2.43 | .1024 | .0490 | .0096 | 2.96 | 2.37 | .0970 | .0458 | .0084 |
| | 2 | 3.19 | 2.61 | .1176 | .0610 | .0144 | 3.02 | 2.39 | .1004 | .0478 | .0092 |
| | 3 | 3.18 | 2.55 | .1142 | .0586 | .0134 | 3.07 | 2.42 | .1010 | .0500 | .0100 |
| | | Group Interaction, $\rho = .5$ | | | | | | | | | |
| 50 | 1 | 3.78 | 2.91 | .1756 | .1022 | .0232 | 3.30 | 2.40 | .1240 | .0578 | .0068 |
| | 2 | 4.66 | 3.65 | .2678 | .1692 | .0566 | 3.40 | 2.37 | .1238 | .0542 | .0056 |
| | 3 | 4.19 | 3.27 | .2174 | .1324 | .0408 | 3.38 | 2.44 | .1278 | .0606 | .0078 |
| 100 | 1 | 3.47 | 2.79 | .1458 | .0822 | .0192 | 3.21 | 2.48 | .1182 | .0546 | .0106 |
| | 2 | 4.09 | 3.29 | .2066 | .1226 | .0386 | 3.28 | 2.40 | .1162 | .0536 | .0074 |
| | 3 | 3.68 | 3.04 | .1612 | .0944 | .0276 | 3.20 | 2.45 | .1104 | .0550 | .0092 |
| 200 | 1 | 3.26 | 2.66 | .1240 | .0652 | .0156 | 3.13 | 2.50 | .1110 | .0548 | .0114 |
| | 2 | 3.61 | 2.93 | .1592 | .0928 | .0234 | 3.19 | 2.42 | .1144 | .0564 | .0068 |
| | 3 | 3.44 | 2.81 | .1382 | .0756 | .0196 | 3.19 | 2.51 | .1144 | .0542 | .0104 |
| 500 | 1 | 3.10 | 2.53 | .1092 | .0578 | .0118 | 3.05 | 2.47 | .1042 | .0526 | .0102 |
| | 2 | 3.22 | 2.65 | .1222 | .0650 | .0150 | 3.04 | 2.42 | .1052 | .0520 | .0080 |
| | 3 | 3.15 | 2.57 | .1136 | .0570 | .0116 | 3.04 | 2.43 | .1036 | .0496 | .0100 |

**Note**: for dgp, 1=normal, 2=normal mixture, and 3=lognormal.

**Table 2b** Empirical Mean, sd and Size of $T_{\texttt{AQS}}^{\texttt{DPPD}}$ for Testing $\lambda = 0$, Heterokedastic Errors.

| | | AQS Test | | | | | SAQS Test | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | dgp | mean | sd | 10% | 5% | 1% | mean | sd | 10% | 5% | 1% |
| | | Group Interaction, $\rho = 0$ | | | | | | | | | |
| 50 | 1 | 4.43 | 3.49 | .2418 | .1550 | .0534 | 3.41 | 2.41 | .1256 | .0550 | .0070 |
| | 2 | 5.49 | 4.35 | .3282 | .2340 | .0986 | 3.56 | 2.38 | .1306 | .0570 | .0076 |
| | 3 | 4.84 | 3.78 | .2830 | .1934 | .0692 | 3.42 | 2.36 | .1244 | .0556 | .0048 |
| 100 | 1 | 3.80 | 3.11 | .1676 | .0966 | .0306 | 3.26 | 2.46 | .1120 | .0546 | .0114 |
| | 2 | 4.56 | 3.87 | .2512 | .1638 | .0604 | 3.34 | 2.47 | .1234 | .0574 | .0096 |
| | 3 | 4.11 | 3.36 | .2076 | .1238 | .0410 | 3.29 | 2.41 | .1160 | .0520 | .0078 |
| 200 | 1 | 3.36 | 2.70 | .1328 | .0698 | .0170 | 3.11 | 2.41 | .1052 | .0506 | .0088 |
| | 2 | 3.97 | 3.33 | .1940 | .1184 | .0362 | 3.32 | 2.53 | .1242 | .0592 | .0124 |
| | 3 | 3.58 | 2.89 | .1566 | .0852 | .0226 | 3.16 | 2.42 | .1104 | .0492 | .0084 |
| 500 | 1 | 3.17 | 2.61 | .1168 | .0582 | .0154 | 3.08 | 2.49 | .1060 | .0534 | .0124 |
| | 2 | 3.40 | 2.76 | .1346 | .0736 | .0178 | 3.12 | 2.40 | .1038 | .0514 | .0080 |
| | 3 | 3.30 | 2.70 | .1334 | .0698 | .0176 | 3.12 | 2.48 | .1128 | .0526 | .0116 |
| | | Rook Contiguity, $\rho = 0$ | | | | | | | | | |
| 50 | 1 | 4.30 | 3.45 | .2258 | .1362 | .0468 | 3.23 | 2.33 | .1044 | .0454 | .0066 |
| | 2 | 5.36 | 4.26 | .3252 | .2246 | .0970 | 3.40 | 2.28 | .1154 | .0468 | .0056 |
| | 3 | 4.87 | 3.84 | .2840 | .1826 | .0716 | 3.39 | 2.32 | .1116 | .0500 | .0066 |
| 100 | 1 | 3.46 | 2.81 | .1378 | .0778 | .0188 | 3.07 | 2.36 | .0956 | .0476 | .0068 |
| | 2 | 4.18 | 3.40 | .2180 | .1300 | .0448 | 3.19 | 2.33 | .1050 | .0508 | .0052 |
| | 3 | 3.79 | 3.08 | .1722 | .1032 | .0318 | 3.13 | 2.35 | .1072 | .0484 | .0074 |
| 200 | 1 | 3.28 | 2.70 | .1212 | .0674 | .0160 | 3.04 | 2.40 | .0992 | .0504 | .0088 |
| | 2 | 3.75 | 3.16 | .1732 | .0982 | .0328 | 3.13 | 2.42 | .1062 | .0514 | .0088 |
| | 3 | 3.52 | 2.88 | .1498 | .0880 | .0196 | 3.12 | 2.41 | .1070 | .0546 | .0080 |
| 500 | 1 | 3.15 | 2.63 | .1158 | .0578 | .0150 | 3.05 | 2.51 | .1064 | .0518 | .0122 |
| | 2 | 3.29 | 2.76 | .1262 | .0688 | .0196 | 3.03 | 2.43 | .0984 | .0474 | .0104 |
| | 3 | 3.22 | 2.62 | .1174 | .0636 | .0150 | 3.04 | 2.40 | .0988 | .0492 | .0106 |
| | | Group Interaction, $\rho = .5$ | | | | | | | | | |
| 50 | 1 | 4.34 | 3.42 | .2298 | .1440 | .0468 | 3.35 | 2.34 | .1168 | .0496 | .0056 |
| | 2 | 5.37 | 4.26 | .3220 | .2214 | .0958 | 3.52 | 2.37 | .1260 | .0568 | .0078 |
| | 3 | 4.78 | 3.79 | .2790 | .1830 | .0656 | 3.42 | 2.35 | .1258 | .0504 | .0048 |
| 100 | 1 | 3.84 | 3.14 | .1770 | .1050 | .0310 | 3.29 | 2.50 | .1192 | .0568 | .0100 |
| | 2 | 4.50 | 3.76 | .2364 | .1502 | .0548 | 3.32 | 2.38 | .1134 | .0522 | .0076 |
| | 3 | 4.07 | 3.29 | .2014 | .1252 | .0334 | 3.25 | 2.34 | .1120 | .0492 | .0060 |
| 200 | 1 | 3.38 | 2.74 | .1378 | .0760 | .0148 | 3.15 | 2.46 | .1144 | .0514 | .0102 |
| | 2 | 3.81 | 3.13 | .1730 | .1020 | .0328 | 3.20 | 2.43 | .1064 | .0524 | .0110 |
| | 3 | 3.64 | 2.93 | .1612 | .0932 | .0244 | 3.22 | 2.48 | .1158 | .0544 | .0110 |
| 500 | 1 | 3.16 | 2.58 | .1164 | .0616 | .0130 | 3.07 | 2.46 | .1062 | .0556 | .0096 |
| | 2 | 3.41 | 2.77 | .1392 | .0724 | .0182 | 3.12 | 2.43 | .1110 | .0492 | .0088 |
| | 3 | 3.28 | 2.63 | .1264 | .0632 | .0146 | 3.10 | 2.42 | .1100 | .0524 | .0070 |

**Note**: for dgp, 1=normal, 2=normal mixture, and 3=lognormal.
Heteroskedasticity $h_{ni} \propto |\frac{1}{T-1}\sum_{t=2}^{T}\Delta X_{it}|$.

**Table 3** Empirical Mean, sd and Size of $T_{\texttt{AQS}}^{\texttt{SDPD4}}$ for Testing $\lambda_1 = \lambda_2 = 0$

| | | AQS Test | | | | | SAQS Test | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | dgp | mean | sd | 10% | 5% | 1% | mean | sd | 10% | 5% | 1% |
| | | Group Interaction, $(\rho, \lambda_3) = (0, 0)$ | | | | | | | | | |
| 50 | 1 | 2.52 | 2.39 | .1592 | .0914 | .0206 | 2.24 | 2.04 | .1198 | .0582 | .0098 |
| | 2 | 3.12 | 2.99 | .2316 | .1430 | .0504 | 2.34 | 2.04 | .1354 | .0640 | .0094 |
| | 3 | 2.73 | 2.66 | .1822 | .1102 | .0296 | 2.22 | 2.00 | .1188 | .0572 | .0092 |
| 100 | 1 | 2.33 | 2.25 | .1354 | .0688 | .0150 | 2.18 | 2.04 | .1160 | .0554 | .0090 |
| | 2 | 2.71 | 2.73 | .1828 | .1098 | .0352 | 2.23 | 2.07 | .1208 | .0610 | .0108 |
| | 3 | 2.48 | 2.52 | .1554 | .0896 | .0262 | 2.19 | 2.09 | .1192 | .0594 | .0122 |
| 200 | 1 | 2.16 | 2.17 | .1158 | .0618 | .0146 | 2.08 | 2.06 | .1086 | .0532 | .0120 |
| | 2 | 2.39 | 2.38 | .1430 | .0834 | .0228 | 2.14 | 2.03 | .1156 | .0578 | .0104 |
| | 3 | 2.25 | 2.21 | .1274 | .0682 | .0150 | 2.11 | 2.01 | .1104 | .0536 | .0090 |
| 500 | 1 | 2.07 | 2.09 | .1050 | .0560 | .0136 | 2.04 | 2.04 | .1012 | .0524 | .0118 |
| | 2 | 2.17 | 2.20 | .1182 | .0670 | .0180 | 2.06 | 2.02 | .1070 | .0544 | .0118 |
| | 3 | 2.17 | 2.15 | .1216 | .0642 | .0140 | 2.10 | 2.05 | .1124 | .0564 | .0108 |
| | | Rook Contiguity, $(\rho, \lambda_3) = (0, 0)$ | | | | | | | | | |
| 50 | 1 | 2.35 | 2.33 | .1456 | .0738 | .0214 | 2.07 | 1.96 | .1056 | .0474 | .0086 |
| | 2 | 3.06 | 2.99 | .2238 | .1382 | .0456 | 2.27 | 2.00 | .1240 | .0596 | .0078 |
| | 3 | 2.68 | 2.61 | .1794 | .1056 | .0306 | 2.17 | 1.98 | .1116 | .0554 | .0088 |
| 100 | 1 | 2.19 | 2.15 | .1220 | .0632 | .0136 | 2.05 | 1.95 | .1034 | .0510 | .0090 |
| | 2 | 2.44 | 2.41 | .1532 | .0886 | .0200 | 2.04 | 1.87 | .1032 | .0430 | .0044 |
| | 3 | 2.37 | 2.35 | .1440 | .0846 | .0216 | 2.10 | 1.98 | .1066 | .0532 | .0086 |
| 200 | 11 | 2.06 | 2.06 | .1168 | .0584 | .0106 | 1.99 | 1.95 | .1084 | .0498 | .0066 |
| | 2 | 2.24 | 2.26 | .1304 | .0686 | .0184 | 2.02 | 1.97 | .1012 | .0448 | .0084 |
| | 3 | 2.12 | 2.10 | .1120 | .0606 | .0114 | 1.98 | 1.90 | .0946 | .0470 | .0066 |
| 500 | 1 | 2.00 | 2.04 | .1006 | .0500 | .0110 | 1.97 | 1.99 | .0964 | .0464 | .0096 |
| | 2 | 2.13 | 2.13 | .1148 | .0566 | .0134 | 2.02 | 1.97 | .1026 | .0478 | .0084 |
| | 3 | 2.01 | 2.02 | .1016 | .0546 | .0114 | 1.95 | 1.93 | .0952 | .0476 | .0088 |
| | | Group Interaction, $(\rho, \lambda_3) = (.5, .3)$ | | | | | | | | | |
| 500 | 1 | 2.55 | 2.45 | .1642 | .0926 | .0252 | 2.25 | 2.05 | .1270 | .0622 | .0098 |
| | 2 | 3.29 | 3.15 | .2474 | .1628 | .0608 | 2.46 | 2.14 | .1484 | .0714 | .0116 |
| | 3 | 2.85 | 2.73 | .1994 | .1224 | .0358 | 2.34 | 2.09 | .1332 | .0664 | .0106 |
| 500 | 1 | 2.30 | 2.28 | .1360 | .0752 | .0180 | 2.14 | 2.06 | .1172 | .0576 | .0108 |
| | 2 | 2.71 | 2.72 | .1808 | .1124 | .0348 | 2.22 | 2.05 | .1266 | .0594 | .0102 |
| | 3 | 2.40 | 2.42 | .1492 | .0858 | .0244 | 2.12 | 2.03 | .1120 | .0582 | .0100 |
| 500 | 1 | 2.19 | 2.19 | .1266 | .0626 | .0146 | 2.11 | 2.08 | .1158 | .0548 | .0110 |
| | 2 | 2.43 | 2.41 | .1496 | .0836 | .0208 | 2.17 | 2.04 | .1188 | .0564 | .0104 |
| | 3 | 2.26 | 2.28 | .1294 | .0726 | .0178 | 2.10 | 2.05 | .1090 | .0558 | .0108 |
| 500 | 1 | 1.97 | 1.96 | .0942 | .0468 | .0086 | 1.94 | 1.92 | .0912 | .0442 | .0086 |
| | 2 | 2.14 | 2.14 | .1138 | .0610 | .0156 | 2.03 | 1.99 | .1008 | .0502 | .0100 |
| | 3 | 2.11 | 2.18 | .1096 | .0572 | .0144 | 2.05 | 2.08 | .0998 | .0524 | .0126 |

**Note**: for dgp, 1=normal, 2=normal mixture, and 3=lognormal.

**Table 4** Empirical Mean, sd and Size of $T_{\mathtt{AQS}}^{\mathtt{SDPD5}}$ for Testing $\lambda_2 = \lambda_3 = 0$

| | | AQS Test | | | | | SAQS Test | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | dgp | mean | sd | 10% | 5% | 1% | mean | sd | 10% | 5% | 1% |
| | | Group Interaction, $(\rho, \lambda_1) = (0, 0)$ | | | | | | | | | |
| 50 | 1 | 2.62 | 2.51 | .1772 | .0990 | .0260 | 2.32 | 2.13 | .1344 | .0690 | .0124 |
| | 2 | 3.23 | 3.13 | .2488 | .1574 | .0530 | 2.43 | 2.12 | .1448 | .0694 | .0116 |
| | 3 | 2.89 | 2.72 | .2018 | .1240 | .0378 | 2.37 | 2.09 | .1366 | .0676 | .0096 |
| 100 | 1 | 2.31 | 2.26 | .1328 | .0740 | .0152 | 2.16 | 2.04 | .1128 | .0564 | .0100 |
| | 2 | 2.74 | 2.73 | .1912 | .1120 | .0328 | 2.29 | 2.11 | .1306 | .0628 | .0104 |
| | 3 | 2.55 | 2.63 | .1658 | .1004 | .0298 | 2.25 | 2.19 | .1322 | .0656 | .0140 |
| 200 | 1 | 2.19 | 2.10 | .1212 | .0616 | .0130 | 2.12 | 1.99 | .1120 | .0548 | .0098 |
| | 2 | 2.51 | 2.59 | .1596 | .0896 | .0270 | 2.24 | 2.16 | .1258 | .0626 | .0120 |
| | 3 | 2.27 | 2.31 | .1334 | .0746 | .0198 | 2.12 | 2.09 | .1182 | .0578 | .0128 |
| 500 | 1 | 2.09 | 2.13 | .1094 | .0606 | .0142 | 2.06 | 2.09 | .1058 | .0578 | .0124 |
| | 2 | 2.16 | 2.13 | .1152 | .0614 | .0130 | 2.05 | 1.97 | .1024 | .0512 | .0088 |
| | 3 | 2.13 | 2.16 | .1132 | .0604 | .0158 | 2.07 | 2.06 | .1066 | .0542 | .0130 |
| | | Rook Contiguity, $(\rho, \lambda_1) = (0, 0)$ | | | | | | | | | |
| 50 | 1 | 2.49 | 2.45 | .1584 | .0914 | .0226 | 2.19 | 2.08 | .1242 | .0586 | .0118 |
| | 2 | 3.17 | 3.07 | .2368 | .1522 | .0504 | 2.39 | 2.09 | .1422 | .0664 | .0098 |
| | 3 | 2.78 | 2.76 | .1880 | .1160 | .0380 | 2.27 | 2.09 | .1224 | .0650 | .0122 |
| 100 | 1 | 2.20 | 2.20 | .1226 | .0640 | .0142 | 2.06 | 2.00 | .1070 | .0492 | .0084 |
| | 2 | 2.53 | 2.54 | .1550 | .0950 | .0262 | 2.13 | 2.00 | .1124 | .0546 | .0084 |
| | 3 | 2.35 | 2.35 | .1420 | .0804 | .0210 | 2.09 | 1.99 | .1100 | .0528 | .0094 |
| 200 | 1 | 2.03 | 1.97 | .1048 | .0502 | .0088 | 1.96 | 1.88 | .0970 | .0436 | .0060 |
| | 2 | 2.25 | 2.22 | .1306 | .0666 | .0150 | 2.01 | 1.88 | .0996 | .0436 | .0054 |
| | 3 | 2.11 | 2.12 | .1126 | .0562 | .0140 | 1.98 | 1.92 | .0934 | .0442 | .0092 |
| 500 | 1 | 2.05 | 2.04 | .1010 | .0518 | .0126 | 2.02 | 2.00 | .0964 | .0486 | .0108 |
| | 2 | 2.11 | 2.15 | .1152 | .0602 | .0150 | 2.01 | 2.00 | .1016 | .0508 | .0110 |
| | 3 | 2.06 | 2.04 | .1058 | .0548 | .0114 | 2.00 | 1.95 | .0994 | .0494 | .0092 |
| | | Group Interaction, $(\rho, \lambda_1) = (.5, .3)$ | | | | | | | | | |
| 50 | 1 | 2.71 | 2.57 | .1844 | .1088 | .0322 | 2.39 | 2.16 | .1466 | .0736 | .0118 |
| | 2 | 3.34 | 3.19 | .2566 | .1640 | .0596 | 2.51 | 2.16 | .1566 | .0778 | .0130 |
| | 3 | 2.99 | 2.87 | .2166 | .1316 | .0408 | 2.44 | 2.16 | .1494 | .0752 | .0128 |
| 100 | 1 | 2.53 | 2.49 | .1660 | .0948 | .0244 | 2.36 | 2.26 | .1472 | .0764 | .0162 |
| | 2 | 2.93 | 2.95 | .2066 | .1302 | .0462 | 2.42 | 2.22 | .1454 | .0792 | .0154 |
| | 3 | 2.69 | 2.64 | .1814 | .1094 | .0326 | 2.38 | 2.22 | .1468 | .0760 | .0150 |
| 200 | 1 | 2.26 | 2.31 | .1354 | .0750 | .0190 | 2.17 | 2.19 | .1240 | .0664 | .0156 |
| | 2 | 2.53 | 2.54 | .1630 | .0930 | .0242 | 2.26 | 2.15 | .1300 | .0638 | .0132 |
| | 3 | 2.44 | 2.42 | .1474 | .0860 | .0224 | 2.28 | 2.20 | .1272 | .0680 | .0138 |
| 500 | 1 | 2.17 | 2.15 | .1184 | .0646 | .0144 | 2.14 | 2.11 | .1134 | .0608 | .0126 |
| | 2 | 2.24 | 2.24 | .1288 | .0698 | .0174 | 2.13 | 2.07 | .1118 | .0576 | .0122 |
| | 3 | 2.20 | 2.21 | .1226 | .0652 | .0160 | 2.13 | 2.12 | .1146 | .0606 | .0130 |

**Note**: for dgp, 1=normal, 2=normal mixture, and 3=lognormal.

**Table 5** Empirical Mean, sd and Size of $T_{\mathtt{AQS}}^{\mathtt{SDPD2}}$ for Testing $\lambda_2 = 0$

| $n$ | dgp | AQS Test | | | | | SAQS Test | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | mean | sd | 10% | 5% | 1% | mean | sd | 10% | 5% | 1% |
| | | Group Interaction, $(\rho, \lambda_1, \lambda_3) = (0, 0, 0)$ | | | | | | | | | |
| 50 | 1 | 0.01 | 1.14 | .1582 | .0872 | .0204 | 0.01 | 1.08 | .1324 | .0678 | .0130 |
| | 2 | 0.00 | 1.15 | .1590 | .0902 | .0236 | -0.01 | 1.08 | .1372 | .0696 | .0132 |
| | 3 | 0.01 | 1.45 | .2592 | .1820 | .0766 | 0.00 | 1.22 | .1860 | .1078 | .0234 |
| 100 | 1 | 0.02 | 1.07 | .1250 | .0708 | .0164 | 0.02 | 1.04 | .1170 | .0586 | .0124 |
| | 2 | -0.03 | 1.09 | .1314 | .0746 | .0182 | -0.03 | 1.05 | .1206 | .0618 | .0122 |
| | 3 | 0.05 | 1.33 | .2122 | .1366 | .0512 | 0.04 | 1.15 | .1538 | .0862 | .0192 |
| 200 | 1 | -0.03 | 1.05 | .1160 | .0572 | .0146 | -0.03 | 1.03 | .1104 | .0522 | .0124 |
| | 2 | 0.02 | 1.04 | .1182 | .0624 | .0160 | 0.02 | 1.02 | .1108 | .0582 | .0124 |
| | 3 | -0.01 | 1.24 | .1788 | .1110 | .0418 | -0.01 | 1.12 | .1440 | .0782 | .0156 |
| 500 | 1 | -0.02 | 1.01 | .1046 | .0550 | .0114 | -0.02 | 1.00 | .1022 | .0530 | .0100 |
| | 2 | 0.01 | 1.02 | .1084 | .0534 | .0108 | 0.01 | 1.01 | .1058 | .0512 | .0100 |
| | 3 | 0.02 | 1.13 | .1388 | .0836 | .0250 | 0.02 | 1.06 | .1190 | .0632 | .0162 |
| | | Rook Contiguity, $(\rho, \lambda_1, \lambda_3) = (0, 0, 0)$ | | | | | | | | | |
| 50 | 1 | 0.02 | 1.07 | .1252 | .0618 | .0140 | 0.03 | 1.02 | .1054 | .0490 | .0076 |
| | 2 | -0.01 | 1.11 | .1386 | .0802 | .0190 | -0.01 | 1.04 | .1174 | .0592 | .0110 |
| | 3 | 0.02 | 1.42 | .2430 | .1660 | .0674 | 0.03 | 1.18 | .1654 | .0920 | .0248 |
| 100 | 1 | 0.00 | 1.04 | .1166 | .0552 | .0122 | 0.00 | 1.01 | .1036 | .0484 | .0092 |
| | 2 | -0.01 | 1.03 | .1116 | .0558 | .0110 | -0.01 | 0.99 | .0982 | .0452 | .0068 |
| | 3 | -0.02 | 1.27 | .1914 | .1236 | .0438 | -0.02 | 1.11 | .1426 | .0790 | .0152 |
| 200 | 1 | 0.01 | 0.99 | .0998 | .0474 | .0100 | 0.01 | 0.98 | .0952 | .0442 | .0090 |
| | 2 | -0.01 | 1.01 | .1084 | .0512 | .0092 | -0.01 | 0.99 | .1006 | .0454 | .0070 |
| | 3 | -0.03 | 1.18 | .1554 | .0958 | .0342 | -0.03 | 1.07 | .1234 | .0638 | .0150 |
| 500 | 1 | 0.02 | 0.98 | .0906 | .0466 | .0076 | 0.02 | 0.98 | .0880 | .0444 | .0074 |
| | 2 | 0.01 | 0.97 | .0942 | .0448 | .0074 | 0.01 | 0.96 | .0912 | .0430 | .0062 |
| | 3 | 0.00 | 1.09 | .1284 | .0716 | .0194 | 0.00 | 1.03 | .1110 | .0550 | .0104 |
| | | Group Interaction, $(\rho, \lambda_1, \lambda_3) = (.5, .2, .2)$ | | | | | | | | | |
| 50 | 1 | 0.00 | 1.16 | .1632 | .0912 | .0238 | 0.00 | 1.10 | .1418 | .0718 | .0162 |
| | 2 | 0.00 | 1.16 | .1578 | .0942 | .0232 | 0.01 | 1.09 | .1334 | .0742 | .0136 |
| | 3 | 0.05 | 1.45 | .2580 | .1754 | .0684 | 0.03 | 1.21 | .1830 | .1008 | .0232 |
| 100 | 1 | 0.00 | 1.11 | .1380 | .0766 | .0204 | -0.01 | 1.07 | .1270 | .0682 | .0146 |
| | 2 | -0.01 | 1.11 | .1372 | .0762 | .0196 | -0.01 | 1.08 | .1240 | .0654 | .0130 |
| | 3 | 0.05 | 1.36 | .2124 | .1408 | .0576 | 0.03 | 1.18 | .1594 | .0846 | .0228 |
| 200 | 1 | 0.02 | 1.01 | .1028 | .0536 | .0128 | 0.02 | 1.00 | .0976 | .0504 | .0104 |
| | 2 | 0.00 | 1.02 | .1088 | .0540 | .0108 | 0.00 | 1.00 | .1032 | .0488 | .0094 |
| | 3 | 0.01 | 1.25 | .1814 | .1174 | .0418 | 0.01 | 1.13 | .1414 | .0812 | .0186 |
| 500 | 1 | 0.01 | 1.08 | .1128 | .0556 | .0138 | 0.01 | 1.07 | .1102 | .0542 | .0136 |
| | 2 | 0.02 | 1.08 | .1204 | .0656 | .0184 | 0.02 | 1.07 | .1186 | .0628 | .0170 |
| | 3 | 0.00 | 1.39 | .1576 | .0962 | .0308 | 0.00 | 1.31 | .1364 | .0780 | .0220 |

**Note**: for dgp, 1=normal, 2=normal mixture, and 3=lognormal.

**Table 6** Empirical Mean, sd and Size of $T_{\texttt{AQS}}^{\texttt{STPD}}$ for Testing $\rho = 0$

| | | AQS Test | | | | | SAQS Test | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | dgp | mean | sd | 10% | 5% | 1% | mean | sd | 10% | 5% | 1% |
| | | Group Interaction, $\lambda = (0,0,0)'$ | | | | | | | | | |
| 50 | 1 | 0.01 | 1.11 | .1370 | .0740 | .0172 | 0.02 | 1.05 | .1128 | .0592 | .0096 |
| | 2 | 0.00 | 1.29 | .2078 | .1334 | .0476 | 0.01 | 1.11 | .1460 | .0704 | .0124 |
| | 3 | -0.01 | 1.18 | .1682 | .0946 | .0264 | 0.01 | 1.07 | .1228 | .0606 | .0110 |
| 100 | 1 | -0.02 | 1.08 | .1260 | .0670 | .0162 | -0.02 | 1.04 | .1144 | .0582 | .0114 |
| | 2 | -0.04 | 1.19 | .1672 | .0974 | .0284 | -0.03 | 1.08 | .1300 | .0654 | .0152 |
| | 3 | -0.01 | 1.12 | .1470 | .0824 | .0220 | -0.01 | 1.06 | .1264 | .0648 | .0132 |
| 200 | 1 | -0.04 | 1.03 | .1132 | .0566 | .0096 | -0.03 | 1.01 | .1080 | .0512 | .0072 |
| | 2 | -0.02 | 1.10 | .1352 | .0748 | .0194 | -0.01 | 1.04 | .1110 | .0574 | .0118 |
| | 3 | -0.01 | 1.06 | .1138 | .0606 | .0162 | 0.00 | 1.02 | .1046 | .0512 | .0114 |
| 500 | 1 | 0.00 | 1.01 | .1014 | .0538 | .0114 | 0.00 | 1.00 | .0988 | .0520 | .0104 |
| | 2 | -0.03 | 1.04 | .1132 | .0588 | .0138 | -0.03 | 1.01 | .1034 | .0488 | .0104 |
| | 3 | 0.01 | 1.02 | .1040 | .0506 | .0130 | 0.02 | 1.00 | .0980 | .0476 | .0116 |
| | | Rook Contiguity, $\lambda = (0,0,0)'$ | | | | | | | | | |
| 50 | 1 | -0.01 | 1.12 | .1478 | .0824 | .0190 | 0.01 | 1.06 | .1238 | .0606 | .0106 |
| | 2 | -0.06 | 1.32 | .2136 | .1370 | .0494 | -0.04 | 1.14 | .1508 | .0822 | .0140 |
| | 3 | -0.05 | 1.22 | .1794 | .1078 | .0324 | -0.03 | 1.10 | .1430 | .0710 | .0142 |
| 100 | 1 | 0.01 | 1.07 | .1292 | .0738 | .0168 | 0.02 | 1.04 | .1150 | .0632 | .0114 |
| | 2 | -0.03 | 1.18 | .1642 | .0970 | .0304 | -0.02 | 1.07 | .1262 | .0622 | .0118 |
| | 3 | 0.00 | 1.12 | .1426 | .0774 | .0194 | 0.00 | 1.05 | .1216 | .0624 | .0086 |
| 200 | 1 | 0.00 | 1.04 | .1154 | .0618 | .0136 | 0.01 | 1.02 | .1096 | .0566 | .0114 |
| | 2 | -0.03 | 1.10 | .1316 | .0754 | .0224 | -0.02 | 1.03 | .1102 | .0558 | .0124 |
| | 3 | -0.02 | 1.07 | .1244 | .0672 | .0172 | -0.01 | 1.03 | .1102 | .0588 | .0128 |
| 500 | 1 | -0.02 | 1.01 | .1048 | .0522 | .0094 | -0.02 | 1.00 | .1026 | .0506 | .0090 |
| | 2 | 0.00 | 1.06 | .1190 | .0650 | .0148 | 0.00 | 1.03 | .1102 | .0560 | .0098 |
| | 3 | -0.03 | 1.04 | .1174 | .0580 | .0128 | -0.02 | 1.02 | .1116 | .0530 | .0100 |
| | | Group Interaction, $\lambda = (.3,.3,.3)'$ | | | | | | | | | |
| 50 | 1 | 0.03 | 1.14 | .1544 | .0858 | .0220 | 0.03 | 1.07 | .1300 | .0634 | .0110 |
| | 2 | 0.01 | 1.28 | .2030 | .1300 | .0464 | 0.02 | 1.12 | .1504 | .0760 | .0112 |
| | 3 | 0.03 | 1.19 | .1656 | .0984 | .0300 | 0.04 | 1.08 | .1318 | .0650 | .0122 |
| 100 | 1 | 0.02 | 1.06 | .1232 | .0678 | .0140 | 0.02 | 1.02 | .1134 | .0582 | .0102 |
| | 2 | -0.03 | 1.19 | .1662 | .0976 | .0308 | -0.02 | 1.07 | .1300 | .0654 | .0110 |
| | 3 | 0.00 | 1.11 | .1384 | .0770 | .0178 | 0.01 | 1.05 | .1172 | .0560 | .0100 |
| 200 | 1 | 0.02 | 1.04 | .1102 | .0582 | .0126 | 0.02 | 1.02 | .1054 | .0526 | .0100 |
| | 2 | -0.01 | 1.10 | .1306 | .0746 | .0176 | 0.00 | 1.03 | .1112 | .0542 | .0094 |
| | 3 | 0.00 | 1.07 | .1234 | .0660 | .0156 | 0.01 | 1.03 | .1096 | .0548 | .0114 |
| 500 | 1 | -0.01 | 1.01 | .1024 | .0528 | .0112 | -0.01 | 1.01 | .0998 | .0506 | .0098 |
| | 2 | 0.00 | 1.06 | .1202 | .0666 | .0160 | 0.00 | 1.03 | .1124 | .0588 | .0116 |
| | 3 | 0.01 | 1.01 | .1072 | .0508 | .0092 | 0.01 | 1.00 | .1024 | .0460 | .0078 |

**Note**: for dgp, 1=normal, 2=normal mixture, and 3=lognormal.

**Table 7** Empirical Mean, sd and Size of $T^{\mathtt{SPD}}_{\mathtt{AQS}}$ for Testing $\rho = \lambda_2 = 0$

| | | `AQS Test` | | | | | `SAQS Test` | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $n$ | dgp | mean | sd | 10% | 5% | 1% | mean | sd | 10% | 5% | 1% |
| | | `Group Interaction,` $\lambda_1 = 0, \lambda_3 = 0$ | | | | | | | | | |
| 50 | 1 | 2.58 | 2.47 | .1718 | .0976 | .0240 | 2.28 | 2.08 | .1338 | .0644 | .0108 |
| | 2 | 3.28 | 3.15 | .2516 | .1646 | .0588 | 2.39 | 2.06 | .1386 | .0650 | .0086 |
| | 3 | 2.90 | 2.79 | .2014 | .1246 | .0400 | 2.33 | 2.09 | .1310 | .0662 | .0112 |
| 100 | 1 | 2.32 | 2.30 | .1338 | .0708 | .0200 | 2.16 | 2.07 | .1132 | .0546 | .0126 |
| | 2 | 2.78 | 2.79 | .1874 | .1166 | .0362 | 2.23 | 2.04 | .1206 | .0604 | .0094 |
| | 3 | 2.45 | 2.43 | .1520 | .0868 | .0220 | 2.15 | 2.01 | .1138 | .0564 | .0094 |
| 200 | 1 | 2.15 | 2.13 | .1198 | .0624 | .0134 | 2.07 | 2.01 | .1092 | .0532 | .0094 |
| | 2 | 2.45 | 2.46 | .1576 | .0902 | .0212 | 2.16 | 2.04 | .1206 | .0592 | .0092 |
| | 3 | 2.26 | 2.22 | .1296 | .0676 | .0164 | 2.10 | 1.99 | .1086 | .0512 | .0094 |
| 500 | 1 | 2.06 | 2.03 | .1096 | .0556 | .0116 | 2.03 | 1.99 | .1068 | .0528 | .0104 |
| | 2 | 2.20 | 2.18 | .1278 | .0650 | .0168 | 2.07 | 1.99 | .1102 | .0510 | .0104 |
| | 3 | 2.13 | 2.13 | .1130 | .0598 | .0150 | 2.05 | 2.02 | .1040 | .0520 | .0116 |
| | | `Rook Contiguity,` $\lambda_1 = 0, \lambda_3 = 0$ | | | | | | | | | |
| 50 | 1 | 2.39 | 2.33 | .1452 | .0782 | .0186 | 2.11 | 1.97 | .1098 | .0512 | .0076 |
| | 2 | 3.14 | 3.02 | .2334 | .1528 | .0528 | 2.30 | 2.00 | .1290 | .0592 | .0068 |
| | 3 | 2.76 | 2.66 | .1908 | .1124 | .0332 | 2.20 | 1.98 | .1172 | .0568 | .0074 |
| 100 | 1 | 2.18 | 2.16 | .1214 | .0638 | .0140 | 2.03 | 1.95 | .1022 | .0474 | .0066 |
| | 2 | 2.59 | 2.69 | .1660 | .0970 | .0314 | 2.09 | 1.96 | .1036 | .0508 | .0086 |
| | 3 | 2.42 | 2.45 | .1462 | .0854 | .0250 | 2.11 | 2.00 | .1080 | .0594 | .0084 |
| 200 | 1 | 2.06 | 2.08 | .1084 | .0550 | .0118 | 1.98 | 1.97 | .0968 | .0472 | .0104 |
| | 2 | 2.39 | 2.40 | .1404 | .0778 | .0212 | 2.11 | 1.98 | .1070 | .0516 | .0078 |
| | 3 | 2.14 | 2.13 | .1160 | .0612 | .0134 | 1.98 | 1.91 | .0984 | .0464 | .0070 |
| 500 | 1 | 2.06 | 2.00 | .1050 | .0560 | .0140 | 2.03 | 1.96 | .1000 | .0510 | .0110 |
| | 2 | 2.13 | 2.07 | .1260 | .0600 | .0110 | 2.01 | 1.89 | .1070 | .0420 | .0060 |
| | 3 | 2.05 | 2.09 | .0980 | .0530 | .0140 | 1.98 | 1.99 | .0960 | .0440 | .0130 |
| | | `Group Interaction,` $\lambda_1 = 0.3, \lambda_3 = 0.3$ | | | | | | | | | |
| 50 | 1 | 2.46 | 2.35 | .1534 | .0842 | .0206 | 2.17 | 1.98 | .1114 | .0528 | .0096 |
| | 2 | 3.24 | 3.14 | .2456 | .1586 | .0554 | 2.36 | 2.06 | .1362 | .0642 | .0098 |
| | 3 | 2.83 | 2.74 | .1974 | .1152 | .0344 | 2.28 | 2.03 | .1262 | .0584 | .0094 |
| 100 | 1 | 2.29 | 2.20 | .1324 | .0670 | .0142 | 2.12 | 1.98 | .1102 | .0498 | .0096 |
| | 2 | 2.66 | 2.64 | .1758 | .1056 | .0286 | 2.15 | 1.94 | .1064 | .0502 | .0072 |
| | 3 | 2.42 | 2.44 | .1498 | .0846 | .0230 | 2.10 | 2.00 | .1084 | .0536 | .0094 |
| 200 | 1 | 2.18 | 2.19 | .1222 | .0638 | .0148 | 2.10 | 2.07 | .1122 | .0560 | .0112 |
| | 2 | 2.38 | 2.41 | .1460 | .0812 | .0216 | 2.10 | 2.01 | .1114 | .0536 | .0094 |
| | 3 | 2.24 | 2.24 | .1290 | .0694 | .0168 | 2.07 | 2.00 | .1096 | .0532 | .0100 |
| 500 | 1 | 2.02 | 2.06 | .1070 | .0570 | .0120 | 1.99 | 2.01 | .1030 | .0540 | .0110 |
| | 2 | 2.10 | 2.19 | .1120 | .0540 | .0140 | 1.99 | 2.01 | .0940 | .0470 | .0110 |
| | 3 | 2.07 | 2.14 | .1120 | .0610 | .0130 | 2.00 | 2.02 | .1030 | .0510 | .0100 |

**Note**: for dgp, 1=normal, 2=normal mixture, and 3=lognormal.