Joint Tests for Dynamic and Spatial Effects in Short Panels with Fixed Effects and Heteroskedasticity*

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

July 10, 2020

Abstract

Simple and reliable tests are proposed for testing the existence of dynamic and/or spatial effects in fixed-effects panel data models with small T and possibly heteroskedastic errors. The tests are constructed based on the *adjusted quasi scores* (AQS), which correct the conditional quasi scores given the initial differences to account for the effect of *initial values*. To improve the finite sample performance, standardized AQS tests are also derived, which are shown to have much improved finite sample properties. All the proposed tests are robust against nonnormality, but some are not robust against cross-sectional heteroskedasticity (CH). A different type of adjustments are made on the AQS functions, leading to a set of tests that are fully robust against unknown CH. Monte Carlo results show excellent finite sample performance of the standardized versions of the AQS tests.

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

JEL classifications: C12, C18, C21, C23.

1. Introduction

Panel data (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 to the current and 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 spatial dynamic panel data model (SDPD) is more appropriate than a regular PD model, or a regular dynamic panel data (DPD) model, or a static spatial panel data (SPD) model. Thus, it is highly desirable to device simple and reliable tests helping applied researchers to choose the most appropriate model.

*An early version was circulated under the title: *Joint Tests for Dynamic and Spatial Effects in Short Panels with Fixed Effects.* I would like to thank Peter Robinson, Fei Jin, Lung-Fei Lee, James LeSage, the participants of the 15th International Workshop on Spatial Econometrics and Statistics, Orleans, 27-28 May 2016 and the 2nd Econometrics Workshop at the Chinese University of Hong Kong, 29 April 2017, Guest Editors Qi Li and Vasilis Sarafidis, and two anonymous referees, for their helpful comments. I am grateful to Singapore Management University for financial support under Grant C244/MSS14E002.

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

$$\begin{aligned} y_t &= \rho y_{t-1} + \lambda_1 W_1 y_t + \lambda_2 W_2 y_{t-1} + X_t\beta + Z\gamma + \mu + \alpha_t 1_n + u_t, \\ & u_t = \lambda_3 W_3 u_t + v_t, \quad t = 1, 2, \ldots, T, \end{aligned} \tag{1.1}$$

where $y_t = (y_{1t}, y_{2t}, \ldots, y_{nt})'$ and $v_t = (v_{1t}, v_{2t}, \ldots, v_{nt})'$ are $n \times 1$ vectors of response values and idiosyncratic errors at time t, and $\{v_{it}\}$ are independent across i and t with mean zero and possibly cross-sectional heteroskedasticity (CH) of unknown form; the scalar parameter ρ characterizes the dynamic effect, λ_1 the spatial lag (SL) effect, λ_2 the space-time lag (STL) effect, and λ_3 the spatial error (SE) effect; $\{X_t\}$ are $n \times p$ matrices containing values of p time-varying exogenous variables, Z is an $n \times q$ matrix containing the values of q time-invariant exogenous variables; β and γ are the usual regression coefficients; $W_r, r = 1, 2, 3$ are the given $n \times n$ spatial weight matrices; and μ 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 (1.1) is fairly general, embedding several important submodels popular in the literature. As T is fixed and small, the time specific effects $\{\alpha_t\}$ are always treated as fixed effects and are merged into the time-varying regressors X_t. The individual specific effects μ can be treated as fixed effects (FE), random effects (RE) or correlated random effects (CRE). Yang (2018a) present a unified, initial conditions free, M-estimation and inference method for the FE-SDPD model, Li and Yang (2020b) extend this M-estimation and inference strategy to allow for unknown CH in the model, and Li and Yang (2020a) present an M-estimation and inference method for the CRE-SDPD (or RE-SDPD) model.[1]

A question arises naturally: in practical applications, 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 $\delta = (\rho, \lambda_1, \lambda_2, \lambda_3)' = (\rho, \lambda')'$. They may be marginal or joint tests (under null, one or more elements of δ are set to zero and the rest are treated as free parameters), or conditional tests (under the null, one or more element of δ are to zero, given the rest already being set to zero).

In this paper, we tackle this general testing problem by focusing on the SDPD model with small T, fixed effects, and unknown CH. The specific tests of interest are as follows.

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

When $H_0^{\texttt{PD}}$ is not rejected, then one proceeds with the regular panel data model with FE and the decision is clear. When $H_0^{\texttt{PD}}$ is rejected, at least one element of δ is not zero and 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 $H_0^{\texttt{DPD}}$: $\lambda = 0$, *the regular FE dynamic panel data (FE-DPD) model suffices.*

If $H_0^{\texttt{DPD}}$ is not rejected, then the cause of rejecting $H_0^{\texttt{PD}}$ is due to ρ being non-zero and the FE-DPD model is chosen; otherwise, one needs to proceed with the following test.

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

If $H_0^{\texttt{STPD}}$ is not rejected, then the cause of rejecting $H_0^{\texttt{PD}}$ is that at least one element of

[1]Other works on short SDPD models include Elhorst (2010), Su and Yang (2015), Qu, et al. (2016), and Kuersteiner and Prucha (2018). However, most of the research on SDPD models focuses on long panels (with large n and large T), see, e.g., Yang, et al. (2006), Mutl (2006), Yu, et al. (2008), Yu and Lee (2010), Lee and Yu (2010, 2012, 2014); Bai and Li (2015), and Shi and Lee (2017).

λ is not zero. In this case, one may proceed further to identify which element of λ is not zero by carrying out conditional tests on one or two elements of λ, 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 λ is non-zero when ρ is treated as a free parameter, and the marginal tests on λ_r should be carried out, respectively, for $r = 1, 2, 3$:

Marginal test $H_0^{\texttt{SDPD1}}$: $\lambda_1 = 0$, *the FE-SDPD model without* λ_1 *suffices.*
Marginal test $H_0^{\texttt{SDPD2}}$: $\lambda_2 = 0$, *the FE-SDPD model without* λ_2 *suffices.*
Marginal test $H_0^{\texttt{SDPD3}}$: $\lambda_3 = 0$, *the FE-SDPD model without* λ_3 *suffices.*

Note that the marginal test $H_0^{\texttt{SDPD3}}$ is quite interesting as the general model (1.1) reduces to a 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. The marginal test $H_0^{\texttt{SDPD2}}$ is also interesting as the null model becomes a SDPD model with both `SL` and `SE` effects, popular in practical applications. Another pair of joint tests of particular interest are,

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

When $H_0^{\texttt{SDPD4}}$ is true, the general model given in (1.1) reduces to a 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 (1.1) to this reduced model has not been considered. When $H_0^{\texttt{SDPD5}}$ is true, the general model reduces to a 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 is not available. The last test that we would like to highlight is:

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

Under $H_0^{\texttt{SDP}}$, the model reduces to a static spatial panel data model with `SL` and SE (or `SARAR`) effects. QML estimation and inference for this model were given by Lee and Yu (2010), LM tests for the spatial effects are given by Debarsy and Ertur (2010), and LM-type tests robust against unknown CH are given by Baltagi and Yang (2013b).

More conditional tests might be of interest besides the ones discussed after the introduction of $H_0^{\texttt{STPD}}$. By conditional tests we mean tests for certain types of effects, give some other effect(s) are removed from the model. For example, given $H_0^{\texttt{SDPD2}}$ is not rejected, i.e., λ_2 is set to zero, one might be interested in testing further whether $\rho = 0$, i.e., whether the static `SARAR` model suffices; given $H_0^{\texttt{STPD}}$ 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.

However, methods for testing the above hypotheses do not seem to be available, in particular, when T is small. Two related works, GMM gradient tests (Taspinar et al., 2017) and robust LM tests (Bera et al., 2019), require a large panel, concern mostly the parametric misspecifications, and do not allow for unknown CH. In contrast, the literature on statistical tests for spatial regression models or static spatial panel data models is much bigger. See, among others, Anselin et al. (1996), Anselin and Bera (1998), Anselin (2001), Kelejian and Prucha (2001), Yang (2010, 2015, 2018c), Born and Breitung (2011), Baltagi and Yang (2013a,b), Robinson and Rossi (2014, 2015a), Jin and Lee (2015, 2018), Liu and Prucha (2018) for spatial regression models; Baltagi et al. (2003), Baltagi et al. (2007), Debarsy and Ertur (2010), Baltagi and Yang (2013a,b), Robinson and Rossi (2015b), and Xu and Yang (2020) for static panel data models.

In this paper, we propose a general and yet simple method, the *adjusted quasi score* (AQS) method, for constructing test statistics for various hypothesis concerning the SDPD models with fixed-effects, small T and possibly heteroskedastic errors. A score-type test is preferred as it requires only the estimation of the null model. The initial constructions of the tests are based on the unified M-estimation method of Yang (2018a): first adjusting the conditional quasi score functions given the initial differences to achieve unbiasedness and consistency, and then developing a martingale difference representation of the AQS function to give a consistent estimate of the variance-covariance matrix of the AQS functions. The resulting AQS tests are shown to have standard asymptotic null behavior and are free from the specifications of the initial conditions. Further corrections are made on the concentrated AQS functions, giving a set of standardized AQS (SAQS) tests with much better finite sample properties. All the proposed tests are robust against nonnormality. Certain tests are fully robust against unknown CH; the others are not and for this alternative modifications are made by following the M-estimation strategy of Li and Yang (2020b) to give tests that are fully robust against unknown CH. Monte Carlo results show excellent performance of the SAQS tests and full robustness of the last test.

The rest of the paper is organized as follows. Section 2 presents the AQS and standardized AQS tests under homoskedasticity. Section 3 presents the AQS tests fully robust against cross-sectional heteroskedasticity. Section 4 present Monte Carlo results. Section 5 concludes the paper. Some necessary technical details are given in Appendix.

2. Adjusted Quasi Score Tests

In this section, we introduce that AQS and standardized AQS tests under the assumptions that the idiosyncratic errors $\{v_{it}\}$ are independent and identically distributed (iid). We identify that some of these tests are automatically robust against unknown CH due to the fact that the spatial weights matrices have zero diagonal elements.

2.1. The AQS function

The methodology we adopt in constructing tests statistics for testing various hypotheses requires the estimation of the null models. In certain cases, e.g., $H_0^{\texttt{PD}}$, the null models are very simple, but in other cases they are not as the null models may still contain the dynamic parameter ρ and/or some of the spatial parameters. Also, the construction of the AQS tests requires the AQS function for the full model. Thus, it is necessary to outline the unified M-estimation method of Yang (2018a). As the current paper focuses on the fixed effects model with small T, the time specific effects are absorbed into the time-varying regressors X_t. First-differencing Model (1.1) to eliminate μ, we have,

$$\Delta y_t = \rho \Delta y_{t-1} + \lambda_1 W_1 \Delta y_t + \lambda_2 W_2 \Delta y_{t-1} + \Delta X_t \beta + \Delta u_t, \;\; \Delta u_t = \lambda_3 W_3 \Delta u_t + \Delta v_t, \quad (2.1)$$

for $t = 2, 3, \ldots, T$. The parameters left in Model (2.1) are $\psi = \{\beta', \sigma_v^2, \rho, \lambda'\}'$. Note that Δy_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 hence Δy_1 is not. Letting ψ_0 be the true value of ψ and $\mathrm{E}(\cdot)$ correspond to ψ_0, Yang's (2018a) M-estimation strategy goes as follows: formulate the conditional quasi likelihood function **as if** Δy_1 is exogenous to give the conditional quasi score vector $S(\psi)$, then adjust $S(\psi)$ to give the AQS vector $S^*(\psi_0) = S(\psi_0) - \mathrm{E}[S(\psi_0)]$, and then estimate ψ by solving the AQS equations $S^*(\psi) = 0$.[2] Some details follow.

[2] Interestingly, this method finds root in Neyman and Scott (1948) on modified likelihood equations. Chudik and Pesaran (2017) use similar ideas to give a bias-corrected method of moments estimation.

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 $\Delta v = \{\Delta v_2', \ldots, \Delta v_T'\}$. Let $\mathbf{W}_r = I_{T-1} \otimes W_r, r = 1, 2, 3$; $B_r(\lambda_r) = I_n - \lambda_r W_r$ and $\mathbf{B}_r(\lambda_r) = I_{T-1} \otimes B_r(\lambda_r)$, for $r = 1$ and 3; and $B_2(\rho, \lambda_2) = \rho I_n + \lambda_2 W_2$ and $\mathbf{B}_2(\rho, \lambda_2) = I_{T-1} \otimes B_2(\rho, \lambda_2)$, where $\otimes$ denotes the Kronecker product and I_m an $m \times m$ identity matrix. Denote $B_1 = B_1(\lambda_1)$ and $B_{10} = B_1(\lambda_{10})$, etc. Assume (i) the errors $\{v_{it}\}$ are iid across i and $t > 0$, (ii) the regressors $\{X_t\}$ are exogenous with respect to $\{v_{it}\}$, (iii) both B_{10}^{-1} and B_{30}^{-1} exist; and (iv) the following 'knowledge' about the process in the past:

Assumption A. *Under Model* (1.1), *(i) the processes started m periods before the start of data collection, the 0th period, and (ii) if $m \geq 1$, Δy_0 is independent of future errors $\{v_t, t \geq 1\}$; if $m = 0$, y_0 is independent of future errors $\{v_t, t \geq 1\}$.*

Yang (2018a) shows: $\mathrm{E}(\Delta Y_{-1} \Delta v') = -\sigma_{v0}^2 \mathbf{D}_{-10} \mathbf{B}_{30}^{-1}$ and $\mathrm{E}(\Delta Y \Delta v') = -\sigma_{v0}^2 \mathbf{D}_0 \mathbf{B}_{30}^{-1}$, where

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

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

and $\mathcal{B} \equiv \mathcal{B}(\rho, \lambda_1, \lambda_2) = B_1^{-1}(\lambda_1) B_2(\rho, \lambda_2)$. These immediately lead to $\mathrm{E}[S(\psi_0)]$, the AQS vector at ψ_0: $S^*(\psi_0) = S(\psi_0) - \mathrm{E}[S(\psi_0)]$, and the AQS vector at a general ψ:

$$S^*(\psi) = \begin{cases} \frac{1}{\sigma_v^2} \Delta X' \Omega^{-1} \Delta u(\theta), \\ \frac{1}{2\sigma_v^4} \Delta u(\theta)' \Omega^{-1} \Delta u(\theta) - \frac{N}{2\sigma_v^2}, \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} \Delta Y_{-1} + \mathrm{tr}(\mathbf{C}^{-1} \mathbf{D}_{-1}), \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} \mathbf{W}_1 \Delta Y + \mathrm{tr}(\mathbf{C}^{-1} \mathbf{D} \mathbf{W}_1), \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' \Omega^{-1} \mathbf{W}_2 \Delta Y_{-1} + \mathrm{tr}(\mathbf{C}^{-1} \mathbf{D}_{-1} \mathbf{W}_2), \\ \frac{1}{\sigma_v^2} \Delta u(\theta)' (C^{-1} \otimes \mathcal{A}) \Delta u(\theta) - (T-1)\mathrm{tr}(G_3), \end{cases} \tag{2.2}$$

where $\theta = (\beta', \rho, \lambda_1, \lambda_2)'$, $\Delta u(\theta) = \mathbf{B}_1(\lambda_1) \Delta Y - \mathbf{B}_2(\rho, \lambda_2) \Delta Y_{-1} - \Delta X \beta$, $G_3 = W_3 B_3^{-1}$, $\mathcal{A} = \frac{1}{2}(W_3' B_3 + B_3' W_3)$, $\Omega = C \otimes (B_3' B_3)^{-1}$, noting $B_3 = B_3(\lambda_3)$, and

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

Solving $S^*(\psi) = 0$ leads to the M-estimator $\hat{\psi}_{\mathtt{M}}$ of ψ. This root-finding process can be simplified by first solving concentrated AQS equations, $S_{\mathtt{c}}^*(\delta) = 0$, with β and σ_v^2 being concentrated out from (2.2), to give the M-estimator $\hat{\delta}_{\mathtt{M}}$ of δ, where

$$S_{\mathtt{c}}^*(\delta) = \begin{cases} \frac{1}{\hat{\sigma}_v^2(\delta)} \Delta \hat{u}(\delta)' \Omega^{-1} \Delta Y_{-1} + \mathrm{tr}(\mathbf{C}^{-1} \mathbf{D}_{-1}), \\ \frac{1}{\hat{\sigma}_v^2(\delta)} \Delta \hat{u}(\delta)' \Omega^{-1} \mathbf{W}_1 \Delta Y + \mathrm{tr}(\mathbf{C}^{-1} \mathbf{D} \mathbf{W}_1), \\ \frac{1}{\hat{\sigma}_v^2(\delta)} \Delta \hat{u}(\delta)' \Omega^{-1} \mathbf{W}_2 \Delta Y_{-1} + \mathrm{tr}(\mathbf{C}^{-1} \mathbf{D}_{-1} \mathbf{W}_2), \\ \frac{1}{\hat{\sigma}_v^2(\delta)} \Delta \hat{u}(\delta)' (C^{-1} \otimes \mathcal{A}) \Delta \hat{u}(\delta) - (T-1)\mathrm{tr}(G_3), \end{cases} \tag{2.3}$$

$\Delta\hat{u}(\delta) = \Delta u(\hat{\beta}(\delta), \rho, \lambda_1, \lambda_2)$, $\hat{\beta}(\delta) = (\Delta X'\Omega^{-1}\Delta X)^{-1}\Delta X'\Omega^{-1}(\mathbf{B}_1\Delta Y - \mathbf{B}_2\Delta Y_{-1})$, and $\hat{\sigma}_v^2(\delta) = \frac{1}{N}\Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta)$. The M-estimators of β and σ_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 (2018a) 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, where $N = n(T-1)$. The M-estimators under the constraints imposed by various hypotheses will remain to be $\sqrt{N}$-consistent and asymptotically normal. It is important to note that the adjustments (non-stochastic terms in (2.2)) are free from the initial conditions, and hence the resulting AQS function and the M-estimators are free from the initial conditions.

2.2. Construction of the AQS test

The AQS functions given in (2.2) 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 some details for the tests defined in Sec. 1. Let $\mathtt{diag}(A)$ form a diagonal matrix by the diagonal elements of a square matrix A and $\mathtt{blkdiag}(A_k)$ form a block-diagonal matrix by matrices $\{A_k\}$. The subscript '$_n$' is often dropped shall no confusion arise.

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 (2.2), i.e., $\Gamma^*(\psi_0) = \frac{1}{N}\mathrm{Var}[S^*(\psi_0)]$. The dynamic nature of Model (1.1) 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 (2018a) 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 same feature of the AQS functions lead to the AQS tests that are free from the initial conditions.

Yang (2018a) developed the representations: $\Delta Y = \mathbb{R}\,\Delta\mathbf{y}_1 + \boldsymbol{\eta} + \mathbb{S}\Delta v$ and $\Delta Y_{-1} = \mathbb{R}_{-1}\Delta\mathbf{y}_1 + \boldsymbol{\eta}_{-1} + \mathbb{S}_{-1}\Delta v$, leading to the expression for the AQS vector at ψ_0 as:

$$S^*(\psi_0) = \begin{cases} \Pi_1'\Delta v, \\ \Delta v'\Phi_1\Delta v - \frac{N}{2\sigma_{v0}^2}, \\ \Delta v'\Psi_1\Delta\mathbf{y}_1 + \Delta v'\Pi_2 + \Delta v'\Phi_2\Delta v + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_{-10}), \\ \Delta v'\Psi_2\Delta\mathbf{y}_1 + \Delta v'\Pi_3 + \Delta v'\Phi_3\Delta v + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_0\mathbf{W}_1), \\ \Delta v'\Psi_3\Delta\mathbf{y}_1 + \Delta v'\Pi_4 + \Delta v'\Phi_4\Delta v + \mathrm{tr}(\mathbf{C}^{-1}\mathbf{D}_{-10}\mathbf{W}_2), \\ \Delta v'\Phi_5\Delta v - (T-1)\mathrm{tr}(G_{30}), \end{cases} \tag{2.4}$$

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 $\mathbb{C}_b = C^{-1}\otimes B_{30}$. Furthermore, $\Delta\mathbf{y}_1 = 1_{T-1}\otimes\Delta y_1$, $\mathbb{R} = \mathtt{blkdiag}(\mathcal{B}_0, \mathcal{B}_0^2, \ldots, \mathcal{B}_0^{T-1})$, $\mathbb{R}_{-1} = \mathtt{blkdiag}(I_n, \mathcal{B}_0, \ldots, \mathcal{B}_0^{T-2})$, $\boldsymbol{\eta} = \mathbb{B}\mathbf{B}_{10}^{-1}\Delta X\beta_0$, $\boldsymbol{\eta}_{-1} = \mathbb{B}_{-1}\mathbf{B}_{10}^{-1}\Delta X\beta_0$, $\mathbb{S} = \mathbb{B}\mathbf{B}_{10}^{-1}\mathbf{B}_{30}^{-1}$, $\mathbb{S}_{-1} = \mathbb{B}_{-1}\mathbf{B}_{10}^{-1}\mathbf{B}_{30}^{-1}$, and

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

The expression (2.4) is the key to the proof of asymptotic normality of $\frac{1}{\sqrt{N}}S^*(\psi_0)$, and to the development of OPMD estimate of the VC matrix of $S^*(\psi_0)$, so that 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 Π, Φ and Ψ are nonstochastic matrices (depending on ψ_0) with Π being $N \times p$ or $N \times 1$, and Φ and Ψ being $N \times N$. As noted in Yang (2018a), 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 Δy_1, which is unavailable. Yang (2018a) 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 Π_t, Φ_{ts} and Ψ_{ts} the submatrices of Π, Φ and Ψ partitioned according to $t, s = 2, \ldots, T$. Define $\Psi_{t+} = \sum_{s=2}^T \Psi_{ts}$, $\Theta = \Psi_{2+}(B_{30}B_{10})^{-1}$, $\Delta y_1^\circ = B_{30}B_{10}\Delta y_1$, and $\Delta y_{1t}^* = \Psi_{t+}\Delta y_1$. Define

$$g_{1i} = \sum_{t=2}^T \Pi_{it}'\Delta v_{it}, \tag{2.5}$$

$$g_{2i} = \sum_{t=2}^T (\Delta v_{it}\Delta\xi_{it} + \Delta v_{it}\Delta v_{it}^* - \sigma_{v0}^2 d_{it}), \tag{2.6}$$

$$g_{3i} = \Delta v_{2i}\Delta\zeta_i + \Theta_{ii}(\Delta v_{2i}\Delta y_{1i}^\circ + \sigma_{v0}^2) + \sum_{t=3}^T \Delta v_{it}\Delta y_{1it}^*, \tag{2.7}$$

where for (2.6), $\xi_t = \sum_{s=2}^T(\Phi_{st}^{u\prime} + \Phi_{ts}^l)\Delta v_s$, $\Delta v_t^* = \sum_{s=2}^T \Phi_{ts}^d \Delta v_s$, and $\{d_{it}\}$ are the diagonal elements of $\mathbf{C}\Phi$; for (2.7), $\{\Delta\zeta_i\} = \Delta\zeta = (\Theta^u + \Theta^l)\Delta y_1^\circ$, and $\mathtt{diag}\{\Theta_{ii}\} = \Theta^d$. Then,

$$\Pi'\Delta v = \sum_{i=1}^n g_{1i}, \tag{2.8}$$

$$\Delta v'\Phi\Delta v - \mathrm{E}(\Delta v'\Phi\Delta v) = \sum_{i=1}^n g_{2i}, \tag{2.9}$$

$$\Delta v'\Psi\Delta \mathbf{y}_1 - \mathrm{E}(\Delta v'\Psi\Delta \mathbf{y}_1) = \sum_{i=1}^n g_{3i}, \tag{2.10}$$

and $\{(g_{1i}', \ g_{2i}, \ g_{3i})', \mathcal{F}_{n,i}\}_{i=1}^n$ form a vector martingale difference (MD) sequence, where $\mathcal{F}_{n,i} = \mathcal{F}_{n,0} \otimes \mathcal{G}_{n,i}$, with $\{\mathcal{G}_{n,i}\}$ being an increasing sequence of σ-fields generated by $(v_{j1}, \ldots, v_{jT}, j = 1, \ldots, i), i = 1, \ldots, n$, and $\mathcal{F}_{n,0}$ the σ-field generated by $(v_0, \Delta y_0)$.

Now, following these results, for each $\Pi_r, r = 1, 2, 3, 4$, defined in (2.4), define g_{1ri} according to (2.5); for each $\Phi_r, r = 1, \ldots, 5$, defined in (2.4), define g_{2ri} according to (2.6); and for each $\Psi_r, r = 1, 2, 3$, defined in (2.4), define g_{3ri} according to (2.7). Define

$$\mathbf{g}_i = (g_{11i}', \ g_{21i}, \ g_{31i} + g_{12i} + g_{22i}, \ g_{32i} + g_{13i} + g_{23i}, \ g_{33i} + g_{14i} + g_{24i}, \ g_{25i})'. \tag{2.11}$$

Then, $S^*(\psi_0) = \sum_{i=1}^n \mathbf{g}_i$, where $\{\mathbf{g}_i, \mathcal{F}_{n,i}\}$ form a vector MD sequence. It follows that $\Gamma^*(\psi_0) = \mathrm{Var}[S^*(\psi_0)] = \sum_{i=1}^n \mathrm{E}(\mathbf{g}_i\mathbf{g}_i')$, and therefore its sample analogue,

$$\widehat{\Gamma}^* = \sum_{i=1}^n \hat{\mathbf{g}}_i\hat{\mathbf{g}}_i', \tag{2.12}$$

gives a consistent OPMD estimator of $\Gamma^*(\psi_0)$, i.e., $\mathrm{plim}_{n\to\infty}\frac{1}{N}\sum_{i=1}^n[\hat{\mathbf{g}}_i\hat{\mathbf{g}}_i' - \Gamma^*(\psi_0)] = 0$, where $\hat{\mathbf{g}}_i$ is obtained by replacing ψ_0 in $\mathbf{g}_i$ by $\hat{\psi}_{\mathtt{M}}$ and Δv by its observed counterpart $\widehat{\Delta} v$, noting that Δy_1 is observed. To construct the AQS tests, the estimates $(\hat{\psi}_{\mathtt{M}}, \widehat{\Delta} v)$ of the full model are replaced by the constrained estimates at the null, $(\tilde{\psi}_{\mathtt{M}}, \widetilde{\Delta} v)$.

To develop the AQS test in a unified manner, let $\delta = (\pi', \varphi')'$ and the null hypothesis specifies $\varphi = 0$. Let $\vartheta = (\beta', \sigma^2, \pi')'$ and therefore $\psi = (\vartheta', \varphi')'$. Let $\Sigma^*(\psi_0) = -\mathrm{E}[\frac{\partial}{\partial\psi'}S^*(\psi_0)]$. Partition $\Sigma^*(\psi)$ and $\Gamma^*(\psi)$ according to ϑ and φ, and denote their submatrices by $\Sigma_{ab}^*(\psi)$ and $\Gamma_{ab}^*(\psi)$, $a = \vartheta, \varphi$, $b = \vartheta, \varphi$. Let $S^*(\psi) = (S_\vartheta^{*\prime}(\psi), S_\varphi^{*\prime}(\psi))'$ and $\mathbf{g}_i = (\mathbf{g}_{i,\vartheta}', \mathbf{g}_{i,\varphi}')'$. Clearly, the construction of the test of $\varphi = 0$ depends on $S_\varphi^*(\tilde{\vartheta}, 0)$ and its variance, where $\tilde{\vartheta}$ is the null estimate of ϑ. Under mild conditions, a Taylor expansion leads to the following asymptotic MD representation:

$$\begin{aligned}\tfrac{1}{\sqrt{N}}S^*_\varphi(\tilde\vartheta,0_k) &= \tfrac{1}{\sqrt{N}}S^*_\varphi(\vartheta_0,0_k) - \tfrac{1}{\sqrt{N}}\Sigma^*_{\varphi\vartheta}\Sigma^{*-1}_{\vartheta\vartheta}S^*_\vartheta(\vartheta_0,0_k) + o_p(1)\\ &= \tfrac{1}{\sqrt{N}}\textstyle\sum_{i=1}^n(\mathbf{g}_{i,\varphi} - \Lambda\mathbf{g}_{i,\vartheta}) + o_p(1),\end{aligned} \tag{2.13}$$

where $\Lambda = \Sigma^*_{\varphi\vartheta}\Sigma^{*-1}_{\vartheta\vartheta}$, and $k = \dim(\varphi)$. Clearly $\{\mathbf{g}_{i,\varphi} - \Lambda\mathbf{g}_{i,\vartheta}\}$ form a vector MD sequence with respect to $\mathcal{F}_{n,i}$. Therefore, $\mathrm{Var}[\frac{1}{\sqrt{N}}S^*_\varphi(\tilde\vartheta,0_k)] = \frac{1}{N}\sum_{i=1}^n[(\mathbf{g}_{i,\varphi}-\Lambda\mathbf{g}_{i,\vartheta})(\mathbf{g}_{i,\varphi}-\Lambda\mathbf{g}_{i,\vartheta})'] + o(1)$. An AQS-based test for testing the hypothesis $H_0: \varphi = 0$ is thus,

$$T_{\mathtt{M}} = S^{*\prime}_\varphi(\tilde\vartheta,0_k)\big\{\textstyle\sum_{i=1}^n(\tilde{\mathbf{g}}_{i,\varphi} - \tilde\Lambda\tilde{\mathbf{g}}_{i,\vartheta})(\tilde{\mathbf{g}}_{i,\varphi} - \tilde\Lambda\tilde{\mathbf{g}}_{i,\vartheta})'\big\}^{-1}S^*_\varphi(\tilde\vartheta,0_k), \tag{2.14}$$

where $\mathtt{M} = \mathtt{PD}$, $\mathtt{DPD}$, $\mathtt{SDPD1}, \cdots, \mathtt{SDPD5}$, and $\mathtt{SPD}$, associated with the null hypotheses defined in Sec. 1, $\tilde\Lambda = \tilde\Sigma^*_{\varphi\vartheta}\tilde\Sigma^{*-1}_{\vartheta\vartheta}$ is the null estimate of Λ^*, and $\tilde{\mathbf{g}}_{i,\vartheta}$ and $\tilde{\mathbf{g}}_{i,\varphi}$ are the null estimates of $\mathbf{g}_{i,\vartheta}$ and $\mathbf{g}_{i,\varphi}$. The asymptotic distribution of $T^{\mathtt{M}}_{\mathtt{AQS}}$, i.e., χ^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 idiosyncratic errors $\{v_{it}\}$ are independent across $i = 1,\ldots,n$ and $t = 0,1,\ldots,T$, with $E(v_{it}) = 0$, $Var(v_{it}) = \sigma^2_{v0}$, and $E|v_{it}|^{4+\epsilon_0} < \infty$ for some $\epsilon_0 > 0$.*

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 ι_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] When homoskedasticity is in question, Assumption B is relaxed to:

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

Additional conditions on the initial differences are necessary when the null model contains the dynamic term, and additioal conditions on B_1 and B_3 are necessary when the null model contains λ_1 and λ_3 terms. These are summarized in the following assumptions.

Assumption E: *For Φ, $n\times n$, uniformly bounded in either row or column sums with elements of uniform order ι_n^{-1}, and ϕ, $n\times 1$, with elements of uniform order $\iota_n^{-1/2}$, (i) $\frac{\iota_n}{n}\Delta y_1'\Phi\Delta y_1 = O_p(1)$, $\frac{\iota_n}{n}\Delta y_1'\Phi\Delta v_2 = O_p(1)$; (ii) $\frac{\iota_n}{n}[\Delta y_1 - \mathrm{E}(\Delta y_1)]'\phi = o_p(1)$; (iii) $\frac{\iota_n}{n}[\Delta y_1'\Phi\Delta y_1 - \mathrm{E}(\Delta y_1'\Phi\Delta y_1)] = o_p(1)$; and (iv) $\frac{\iota_n}{n}[\Delta y_1'\Phi\Delta v_2 - \mathrm{E}(\Delta y_1'\Phi\Delta v_2)] = o_p(1)$.*

Assumption F: *B_1^{-1} and B_2^{-1} exist, and are uniformly bounded in both row and column sums in absolute value, for (λ_1,λ_3) in a neighborhood of $(\lambda_{10},\lambda_{30})$.*

Theorem 2.1. *Under Assumptions A-F, if $\tilde\vartheta$ is $\sqrt{N}$-consistent, we have under $H_0^{\mathtt{M}}$, $T_{\mathtt{M}} \xrightarrow{D} \chi^2_k$, as $n \to \infty$, where $\mathtt{M}$ denotes a null model specified in Sec. 1.*

Note that in a special case where $\Gamma^* \approx \Sigma^*$ at the null, i.e., the information matrix equality (IME) holds (asymptotically), the AQS test is asymptotically equivalent to

[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 (2018a). However, this feature is not explicitly reflected in the subsequent developments as the implementations of the tests do not require ι.

$$T_{\texttt{M},0} = S^{*\prime}(\tilde{\psi})(\textstyle\sum_{i=1}^{n} \tilde{\mathbf{g}}_i \tilde{\mathbf{g}}_i')^{-1} S^*(\tilde{\psi}), \tag{2.15}$$

where $\tilde{\psi} = (\tilde{\vartheta}', 0_k')'$. The cases under which the above can be true are those with the null model being a static panel data model (i.e., $\rho = \lambda_2 = 0$) and the errors are Gaussian.

To facilitate the practical applications of the AQS tests, we now present details for each of the hypothesis postulated in Sec. 1 so that a specific test can directly be applied without going through the complicated general case. More interestingly, we show that certain tests are valid under Assumption B*, i.e., robust against unknown CH.

Joint test $H_0^{\texttt{PD}}$: $\delta = 0$. Under $H_0^{\texttt{PD}}$, the model SDPD(δ) 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 $H_0^{\texttt{PD}}$, $B_1 = B_3 = I_n$, and $B_2 = \mathbf{0}_n$ where $\mathbf{0}_n$ denotes an $n \times n$ matrix of zeros. It is easy to see that $\mathrm{E}[S^*(\psi_0)|_{H_0^{\texttt{PD}}}] = 0$ and that $\tilde{\beta}$ and $\tilde{\sigma}_v^2$ are robust against unknown CH.

Corollary 2.1. *Under Assumptions A, B*, C and D, $T_{\texttt{PD}}|_{H_0^{\texttt{PD}}} \xrightarrow{D} \chi_4^2$, as $n \to \infty$.*

The very attractive feature of this joint test is that it is robust against unknown CH 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 the given 'condition' the model becomes a pure panel data model.

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 β and σ_v^2, given ρ, 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 ρ under $H_0^{\texttt{DPD}}$ is

$$\tilde{\rho} = \arg\left\{ \tfrac{1}{\tilde{\sigma}_v^2(\rho)} \Delta \tilde{v}'(\rho) \mathbf{C}^{-1} \Delta Y_{-1} + n\big(\tfrac{1}{1-\rho} - \tfrac{1-\rho^T}{T(1-\rho)^2}\big) = 0 \right\}, \tag{2.16}$$

leading to the constrained M estimators of β and σ_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 ϑ is thus $\tilde{\vartheta} = (\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho})'$. The following lemma shows that the restricted M-estimator $\tilde{\rho}$ defined in (2.16) is robust against unknown CH.[4]

Lemma 2.1. *Under Assumptions A, B*, and C-E, if ρ_0 is in the interior of a compact parameter space, then for the DPD model, we have, as $n \to \infty$, $\tilde{\vartheta} = (\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho})' \xrightarrow{p} \vartheta_0$, and $\sqrt{N}(\tilde{\vartheta} - \vartheta_0) \xrightarrow{D} N(0, \Psi)$, for a suitably defined Ψ.*

Corollary 2.2. *Under the assumptions of Lemma 2.1, $T_{\texttt{DPD}}|_{H_0^{\texttt{DPD}}} \xrightarrow{D} \chi_3^2$, as $n \to \infty$.*

Corollary 2.2 presents another interesting result: $T_{\texttt{DPD}}$ is robust against both nonnormality and unknown CH, which applies to all tests with a pure DPD model at null.

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

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

[4]The concentrated AQS function for ρ contained in (2.16) 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 (2018a,b) for more discussions.

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 those given below (2.3) by setting $\rho = 0$. Let $\tilde{\beta} = \tilde{\beta}(\tilde{\lambda})$, $\tilde{\sigma}_v^2 = \tilde{\sigma}_v^2(\tilde{\lambda})$, and $\tilde{\vartheta} = \{\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\lambda}'\}'$. Based on the result of Li and Yang (2020b), it is easy to see that $\text{plim}_{n\to\infty}\frac{1}{n}S^*(\psi_0)|_{\rho=0} \neq 0$ under unknown CH. Therefore $\tilde{\vartheta}$ cannot be consistent under unknown CH and $T_{\texttt{STPD}}$ is generally not robust against unknown CH. Sec. 3 presents a CH-robust version of this test.

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` 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:

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

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 those given below (2.3) 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 $\tilde{\psi} = \{\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho}, \tilde{\lambda}_1, 0, \tilde{\lambda}_3\}'$. We obtain the AQS test statistic $T_{\texttt{SPDD2}}$ from (2.14).

Joint test $H_0^{\texttt{SDPD4}}$: $\lambda_1 = \lambda_2 = 0$. This is an interesting 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 with initial observations being modeled. In this case, $B_1 = I_n$ and $B_2 = \rho I_n$, and the constrained M-estimators $\tilde{\rho}$ and $\tilde{\lambda}_3$ solve:

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

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 those given below (2.3) 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 $\tilde{\psi} = \{\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho}, 0, 0, \tilde{\lambda}_3\}'$. We obtain from (2.14) the AQS test $T_{\texttt{SDPD4}}$ for testing $H_0^{\texttt{SDPD4}}$.

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. In this case, $B_2 = \rho I_n$ and $B_3 = I_n$, and the constrained M-estimators $\tilde{\rho}$ and $\tilde{\lambda}_1$ solve:

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

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 $\tilde{\sigma}_v^2(\rho,\lambda_1)$ are those given below (2.3) 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 $\tilde{\psi} = \{\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho}, \tilde{\lambda}_1, 0, 0\}'$. We obtain from (2.14) the AQS test $T_{\texttt{SDPD5}}$ for testing $H_0^{\texttt{SDPD5}}$.

Joint test $H_0^{\texttt{SPD}}$: $\rho = \lambda_2 = 0$. Under the null, $B_2 = 0$ and $\mathbf{D} = -\mathbf{C}\mathbf{B}_1^{-1}$, and the model becomes the static `SARAR` model. The constrained M-estimators $\tilde{\lambda}_1$ and $\tilde{\lambda}_3$ of λ_1 and λ_3 solve the following estimating equations (see also Lee and Yu (2010)):

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

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 those given below (2.3) 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 $\tilde{\psi} =$

$\{\tilde{\beta}', \tilde{\sigma}_v^2, 0, \tilde{\lambda}_1, 0, \tilde{\lambda}_3\}'$. We obtain from (2.14) the AQS test $T_{\texttt{SPD}}$ for testing $H_0^{\texttt{SPD}}$.

Conditional tests are those for testing whether the model can be further reduced, given that it has already been reduced. For example, $H_0^{\texttt{PD1}}$: $\lambda_1 = 0$, given $\lambda_2 = \lambda_3 = 0$; $H_0^{\texttt{PD3}}$: $\lambda_3 = 0$, given $\lambda_1 = \lambda_2 = 0$; $H_0^{\texttt{SPD}}$: $\rho = 0$, given $\lambda_2 = 0$. The last conditional test says that based on the model without λ_2, we want to see if $\rho = 0$, i.e., if the model full SDPD model can be reduced to a regular SPD model. The conditional tests conditional upon $\rho = \lambda_2 = 0$ are the tests of model reduction for the regular SPD 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 unknown CH in light of Corollaries 2.1 and 2.2, and some can be made to be robust against unknown CH in light of Baltagi and Yang (2013b). Given the fact that the OPMD estimator of the VC matrix of AQS functions are robust against unknown CH, any AQS or SAQS test can be made to be CH-robust, provided the AQS function is made so. Instead of discussing this for the individual AQS or SAQS test, a general CH-robut method is given in Sec. 3.

All the tests developed above can be implemented in a unified manner based on the general expressions of the AQS function given in (2.2) or (2.4), and the general OPMD estimate of its VC matrix given in (2.12). $\tilde{\Sigma}^*$ can be $\Sigma^*(\tilde{\psi})$ or $-\frac{\partial}{\partial \psi} S^*(\psi)|_{\psi=\tilde{\psi}_{\texttt{M}}}$. 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.

2.3. 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 β and σ_v^2 are not taken into account when constructing the test statistics. It is thus desirable to find simple ways to improve the finite sample performance of these tests. Clearly, after β_0 and σ_v^2 being replaced by $\hat{\beta}(\delta_0)$ and $\hat{\sigma}_v(\delta_0)$ in the last four components of $S^*(\psi_0)$ given in (2.2), 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).

Rewrite the numerator, $\hat{\sigma}_v^2(\delta) S_{\texttt{c}}^*(\delta)$, of the concentrated AQS function in (2.3) as

$$S_{\texttt{c,N}}^*(\delta) = \begin{cases} \Delta\hat{u}(\delta)'\Omega^{-1}\Delta Y_{-1} + \phi_1 \Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta), \\ \Delta\hat{u}(\delta)'\Omega^{-1}\mathbf{W}_1\Delta Y + \phi_2 \Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta), \\ \Delta\hat{u}(\delta)'\Omega^{-1}\mathbf{W}_2\Delta Y_{-1} + \phi_3 \Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta), \\ \Delta\hat{u}(\delta)'(C^{-1} \otimes \mathcal{A})\Delta\hat{u}(\delta) - \phi_4 \Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta), \end{cases} \tag{2.17}$$

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 $\phi_4 = \frac{1}{n}\text{tr}(G_3)$. The ideas are: finding the mean of $S_{\texttt{c,N}}^*(\delta_0)$ and recentering, and then finding the variance estimate of the recentered $S_{\texttt{c,N}}^*(\delta_0)$ and restandardizing.

Letting $\Omega^{\frac{1}{2}}$ be the symmetric square root matrix of Ω, 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 δ_0, $\Omega_0^{-\frac{1}{2}}(\mathbf{B}_{10}\Delta Y - \mathbf{B}_{20}\Delta Y_{-1}) = \Delta X^*\beta_0 + \Omega_0^{-\frac{1}{2}}\mathbf{B}_{30}^{-1}\Delta v$, we obtain

$$S^*_{\mathtt{c,N}}(\delta_0) = \begin{cases} \Delta v'\mathbf{B}_{30}'^{-1}\mathbf{M}_0^*\Delta Y_{-1} + \phi_{10}\Delta v'\mathbf{M}_0^{**}\Delta v, \\ \Delta v'\mathbf{B}_{30}'^{-1}\mathbf{M}_0^*\mathbf{W}_1\Delta Y + \phi_{20}\Delta v'\mathbf{M}_0^{**}\Delta v, \\ \Delta v'\mathbf{B}_{30}'^{-1}\mathbf{M}_0^*\mathbf{W}_2\Delta Y_{-1} + \phi_{30}\Delta v'\mathbf{M}_0^{**}\Delta v, \\ \Delta v'\mathbf{M}_0^{**}(C\otimes G_{30})\mathbf{M}_0^{**}\Delta v - \phi_{40}\Delta v'\mathbf{M}_0^{**}\Delta v, \end{cases} \tag{2.18}$$

where $\mathbf{M}^* = \Omega^{-\frac{1}{2}}\mathbf{M}\Omega^{-\frac{1}{2}}$ and $\mathbf{M}^{**} = \mathbf{B}_3'^{-1}\mathbf{M}^*\mathbf{B}_3^{-1}$, and thus $\mathrm{E}[S^*_{\mathtt{c,N}}(\delta_0)$ with elements: $\mu_{\rho_0} = \sigma_{v0}^2\mathrm{tr}[(\mathbf{B}_{30}'\mathbf{B}_{30})^{-1}\mathbf{M}_0^*(\phi_{10}\mathbf{C} - \mathbf{D}_{-10})]$, $\mu_{\lambda_{10}} = \sigma_{v0}^2\mathrm{tr}[(\mathbf{B}_{30}'\mathbf{B}_{30})^{-1}\mathbf{M}_0^*(\phi_{20}\mathbf{C} - \mathbf{W}_1\mathbf{D}_0)]$, $\mu_{\lambda_{20}} = \sigma_{v0}^2\mathrm{tr}[(\mathbf{B}_{30}'\mathbf{B}_{30})^{-1}\mathbf{M}_0^*(\phi_{30}\mathbf{C} - \mathbf{W}_2\mathbf{D}_{-10})]$, and $\mu_{\lambda_{30}} = \sigma_{v0}^2\mathrm{tr}[\mathbf{M}_0^{**}(C\otimes G_{30} - \phi_{40}\mathbf{C})]$.[5] Thus, the recentered AQS function takes the form:

$$S^\diamond_{\mathtt{c,N}}(\delta) = S^*_{\mathtt{c,N}}(\delta) - (\mu_\rho, \mu_{\lambda_1}, \mu_{\lambda_2}, \mu_{\lambda_3})'. \tag{2.19}$$

To develop an OPMD estimate of the VC matrix of $S^\diamond_{\mathtt{c,N}}(\delta_0)$, similar to (2.4) we have,

$$S^\diamond_{\mathtt{c,N}}(\delta_0) = \begin{cases} \Delta v'\Psi_1\Delta\mathbf{y}_1 + \Delta v'\Pi_1 + \Delta v'\Phi_1\Delta v - \mu_{\rho_0}, \\ \Delta v'\Psi_2\Delta\mathbf{y}_1 + \Delta v'\Pi_2 + \Delta v'\Phi_2\Delta v - \mu_{\lambda_{10}}, \\ \Delta v'\Psi_3\Delta\mathbf{y}_1 + \Delta v'\Pi_3 + \Delta v'\Phi_3\Delta v - \mu_{\lambda_{20}}, \\ \Delta v'\Phi_4\Delta v - \mu_{\lambda_{30}}, \end{cases} \tag{2.20}$$

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

Similar to $\{\mathbf{g}_i\}$ defined based on (2.4), we define $\{\mathbf{g}_i^\circ\}$ based on (2.20). Now, $\{\mathbf{g}_i^\diamond\}$ are functions of unknown parameters δ_0 and unobserved errors Δv. Replacing δ_0 by $\tilde{\delta}$ and Δv by $\widetilde{\Delta} v$ in $\{\mathbf{g}_i^\diamond\}$ to give $\{\tilde{\mathbf{g}}_i^\diamond\}$, one obtains an OPMD estimate of $\Gamma^\diamond(\delta_0) = \mathrm{Var}[S^\diamond_{\mathtt{c,N}}(\delta_0)]$:

$$\widehat{\Gamma}^\diamond = \textstyle\sum_{i=1}^n \tilde{\mathbf{g}}_i^\diamond\tilde{\mathbf{g}}_i^{\diamond\prime}. \tag{2.21}$$

Again, to develop the standardized AQS tests in a unified manner, recall $\delta = (\pi', \varphi')'$ and the null hypothesis specifies $\varphi = 0$. Let $\Sigma^\diamond(\delta_0) = -\mathrm{E}[\frac{\partial}{\partial\delta}S^\diamond(\delta_0)]$. Partition $\Sigma^\diamond(\delta)$ and $\Gamma^\diamond(\delta)$ according to π and φ, and denote their submatrices by $\Sigma^\diamond_{ab}(\delta)$ and $\Gamma^\diamond_{ab}(\delta)$, $a = \pi, \varphi$, $b = \pi, \varphi$. Let $S^\diamond(\delta) = (S_\pi^{\diamond\prime}(\delta), S_\varphi^{\diamond\prime}(\delta))'$ and $\mathbf{g}_i^\diamond = (\mathbf{g}_{i,\pi}^{\diamond\prime}, \mathbf{g}_{i,\varphi}^{\diamond\prime})'$. Now, the construction of the test of $\varphi = 0$ depends on $S^*_\varphi(\tilde{\pi}, 0)$ and its variance, where $\tilde{\pi}$ is the null estimate of π. Similar to (2.13), a Taylor expansion leads to the following asymptotic MD representation:

$$\begin{aligned} \tfrac{1}{\sqrt{N}}S^\diamond_\varphi(\tilde{\pi}, 0_k) &= \tfrac{1}{\sqrt{N}}S^\diamond_\varphi(\pi_0, 0_k) - \tfrac{1}{\sqrt{N}}\Sigma^\diamond_{\varphi\pi}\Sigma^{\diamond-1}_{\pi\pi}S^\diamond_\pi(\pi_0, 0_k) + o_p(1) \\ &= \tfrac{1}{\sqrt{N}}\textstyle\sum_{i=1}^n(\mathbf{g}^\diamond_{i,\varphi} - \Lambda^\diamond\mathbf{g}^\diamond_{i,\pi}) + o_p(1), \end{aligned} \tag{2.22}$$

where $\Lambda^\diamond = \Sigma^\diamond_{\varphi\pi}\Sigma^{\diamond-1}_{\pi\pi}$. Therefore, the standardized AQS (SAQS) test statistic for testing $H_0: \varphi = 0$ takes a similar form as the AQS test:

$$T^\diamond_{\mathtt{M}} = S^{\diamond\prime}_\varphi(\tilde{\pi}, 0_k)\big\{\textstyle\sum_{i=1}^n(\tilde{\mathbf{g}}^\diamond_{i,\varphi} - \tilde{\Lambda}^\diamond\tilde{\mathbf{g}}^\diamond_{i,\pi})(\tilde{\mathbf{g}}^{\diamond\prime}_{i,\varphi} - \tilde{\Lambda}^\diamond\tilde{\mathbf{g}}^\diamond_{i,\pi})'\big\}^{-1}S^\diamond_\varphi(\tilde{\pi}, 0_k), \tag{2.23}$$

where `M` corresponds to `PD`, `DPD`, `SDPDr`, etc., for testing the hypotheses $H_0^{\mathtt{PD}}$, $H_0^{\mathtt{DPD}}$, $H_0^{\mathtt{SDPDr}}$, etc., postulated in Sec. 1. As in (2.15), if IME holds asymptotically, i.e., $\Sigma^\diamond = \Gamma^\diamond + o(N)$, the test can be simplified to $T^\diamond_{\mathtt{M,0}} = S^{\diamond\prime}(\tilde{\delta})(\sum_{i=1}^n\tilde{\mathbf{g}}_i^\diamond\tilde{\mathbf{g}}_i^{\diamond\prime})^{-1}S^\diamond(\tilde{\delta})$, where $\tilde{\delta} = (\tilde{\pi}', 0_k')'$. Furthermore, if null specifies $\delta = 0$, $T^\diamond_{\mathtt{PD}}$ reduces to $T^\diamond_{\mathtt{PD,0}}$ and there is no need of (2.22).

Theorem 2.2. *Under Assumptions A-D, if $\tilde{\pi}$ is $\sqrt{N}$-consistent, we have under $H_0^{\mathtt{M}}$, $T^{\mathtt{M}}_{\mathtt{SAQS}} \xrightarrow{D} \chi_k^2$, as $n\to\infty$, where* `M` *denotes a null model specified in Sec. 1.*

[5] As $\mathbf{M}^* = \Omega^{-1} - \Omega^{-1}\Delta X(\Delta X'\Omega^{-1}\Delta X)^{-1}\Delta X'\Omega^{-1}$, calculations of $\Omega^{\frac{1}{2}}$ and $\Omega^{-\frac{1}{2}}$ are avoided.

Monte Carlo results show that the SAQS tests can offer much improvements over the AQS tests when n is not large, particularly when spatial dependence is heavy. In each SAQS test, the null estimate $\tilde{\pi}$ can be obtained in the same way as that for the AQS test or solving a subset of equations obtained from $S^{\diamond}_{\mathtt{c,N}}(\delta)$, and $T^{\diamond}_{\mathtt{M}}$ is implemented similarly.

All the conditional AQS tests discussed in Sec. 2.2 have their counterparts based on the standardized AQS function. Similar to the case of the regular AQS tests presented in Sec. 2.2, the standardized AQS tests can also be implemented in a unified manner based on the general expressions (2.19) or (2.20), the VC matrix estimate defined in (2.21), and $\tilde{\Sigma}^{\diamond} = -\frac{\partial}{\partial} S^{\diamond}_{\mathtt{c,N}}(\delta)|_{\delta=\tilde{\delta}_{\mathtt{M}}}$. Similar to the AQS tests $T_{\mathtt{PD}}$ and $T_{\mathtt{DPD}}$, the two standardized AQS tests, $T^{\diamond}_{\mathtt{PD}}$ and $T^{\diamond}_{\mathtt{DPD}}$, are also robust against both nonnormality and unknown CH. Others are in general robust only against nonnormality as the corresponding AQS tests. Therefore, it is desirable to have AQS tests fully robust against unknown CH.

3. CH-Robust AQS Tests

As indicated in the early section, when the null model involves both dynamic and spatial parameters, the AQS tests may not be robust against the unknown CH, and there is no simple way to further adjust the AQS function to make it CH-robust. In this section, we introduce an alternative CH-robust AQS method, to give a set of CH-robust tests.

3.1. The CH-robust M-estimation

Li and Yang (2020b) extend Yang (2018a) to propose CH-robust estimation and inference method for Model (2.1), using an alternative way of adjusting the conditional QS functions to give a set of CH-robust AQS functions:

$$S^{*}_{\mathtt{H}}(\psi) = \begin{cases} \frac{1}{\sigma_v^2}\Delta X'\Omega^{-1}\Delta u(\theta), \\ \frac{1}{2\sigma_v^4}\Delta u(\theta)'\Omega^{-1}\Delta u(\theta) - \frac{N}{2\sigma_v^2}, \\ \frac{1}{\sigma_v^2}\Delta u(\theta)'\Omega^{-1}\Delta Y_{-1} + \frac{1}{\sigma_v^2}\Delta u(\theta)'\mathbf{E}_{\rho}\Delta u(\theta), \\ \frac{1}{\sigma_v^2}\Delta u(\theta)'\Omega^{-1}\mathbf{W}_1\Delta Y + \frac{1}{\sigma_v^2}\Delta u(\theta)'\mathbf{E}_{\lambda_1}\Delta u(\theta), \\ \frac{1}{\sigma_v^2}\Delta u(\theta)'\Omega^{-1}\mathbf{W}_2\Delta Y_{-1} + \frac{1}{\sigma_v^2}\Delta u(\theta)'\mathbf{E}_{\lambda_2}\Delta u(\theta), \\ \frac{1}{\sigma_v^2}\Delta u(\theta)'\left[C^{-1}\otimes(\mathcal{A}-\mathbf{E}_{\lambda_3})\right]\Delta u(\theta), \end{cases} \tag{3.1}$$

where $(\mathbf{E}_{\rho}, \mathbf{E}_{\lambda_1}, \mathbf{E}_{\lambda_2}) = \Omega^{-1}\mathbf{C}^{-1}(\mathbf{D}_{-1}, \mathbf{W}_1\mathbf{D}, \mathbf{W}_2\mathbf{D}_{-1})$, and $\mathbf{E}_{\lambda_3} = B_3'\mathtt{diag}(G_3)[\mathtt{diag}(B_3^{-1})]^{-1}$.

Solving the estimating equations, $S^{*}_{\mathtt{CH}}(\psi) = 0$, gives the CH-robust M-estimator $\hat{\psi}_{\mathtt{H}}$. This can be done by first solving the equations for β and σ_v^2, given $\delta = (\rho, \lambda')'$, to give $\hat{\beta}_{\mathtt{H}}(\delta) = (\Delta X'\Omega^{-1}\Delta X)^{-1}\Delta X'\Omega^{-1}(\mathbf{B}_1\Delta Y - \mathbf{B}_2\Delta Y_{-1})$, and $\hat{\sigma}^2_{v,\mathtt{H}}(\delta) = \frac{1}{N}\Delta\hat{u}(\delta)'\Omega^{-1}\Delta\hat{u}(\delta)$, where $\Delta\hat{u}(\delta) = \Delta u(\hat{\beta}(\delta), \rho, \lambda_1, \lambda_2)$. Then, substituting $\hat{\beta}_{\mathtt{H}}(\delta)$ and $\hat{\sigma}^2_{v,\mathtt{H}}(\delta)$ back into the last four components of (3.1) gives the concentrated AQS functions:

$$S^{*c}_{\mathtt{H}}(\delta) = \begin{cases} \frac{1}{\hat{\sigma}^2_{v,\mathtt{M}}(\delta)}\Delta\hat{u}(\delta)'\Omega^{-1}\Delta Y_{-1} + \frac{1}{\hat{\sigma}^2_{v,\mathtt{M}}(\delta)}\Delta\hat{u}(\delta)'\mathbf{E}_{\rho}\Delta\hat{u}(\delta), \\ \frac{1}{\hat{\sigma}^2_{v,\mathtt{M}}(\delta)}\Delta\hat{u}(\delta)'\Omega^{-1}\mathbf{W}_1\Delta Y + \frac{1}{\hat{\sigma}^2_{v,\mathtt{M}}(\delta)}\Delta\hat{u}(\delta)'\mathbf{E}_{\lambda_1}\Delta\hat{u}(\delta), \\ \frac{1}{\hat{\sigma}^2_{v,\mathtt{M}}(\delta)}\Delta\hat{u}(\delta)'\Omega^{-1}\mathbf{W}_2\Delta Y_{-1} + \frac{1}{\hat{\sigma}^2_{v,\mathtt{M}}(\delta)}\Delta\hat{u}(\delta)'\mathbf{E}_{\lambda_2}\Delta\hat{u}(\delta), \\ \frac{1}{\hat{\sigma}^2_{v,\mathtt{M}}(\delta)}\Delta\hat{u}(\delta)'\left[C^{-1}\otimes(\mathcal{A}-\mathbf{E}_{\lambda_3})\right]\Delta\hat{u}(\delta). \end{cases} \tag{3.2}$$

Solving $S^{*c}_{\mathtt{H}}(\delta) = 0$ gives the CH-robust M-estimator $\hat{\delta}_{\mathtt{H}}$ of δ, and then the CH-robust M-estimators of β and σ_v^2: $\hat{\beta}_{\mathtt{H}} \equiv \hat{\beta}_{\mathtt{H}}(\hat{\delta}_{\mathtt{H}})$ and $\hat{\sigma}^2_{v,\mathtt{H}} \equiv \hat{\sigma}^2_{v,\mathtt{H}}(\hat{\delta}_{\mathtt{H}})$.

3.2. The CH-robust AQS tests

By the representations for ΔY and ΔY_{-1} used in Sec. 2.1 and using the relationship $\Delta u = \mathbf{B}_{30}^{-1}\Delta v$, the AQS function at ψ_0 can be written as

$$S_{\mathtt{H}}^{*}(\psi_0) = \begin{cases} \Pi_1'\Delta v, \\ \Delta v'\Phi_1\Delta v - \frac{n(T-1)}{2\sigma_{v0}^2}, \\ \Delta v'\Psi_1\Delta \mathbf{y}_1 + \Pi_2'\Delta v + \Delta v'\Phi_2\Delta v, \\ \Delta v'\Psi_2\Delta \mathbf{y}_1 + \Pi_3'\Delta v + \Delta v'\Phi_3\Delta v, \\ \Delta v'\Psi_3\Delta \mathbf{y}_1 + \Pi_4'\Delta v + \Delta v'\Phi_4\Delta v, \\ \Delta v'\Phi_5\Delta v, \end{cases} \tag{3.3}$$

where $\Pi_1 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_{b0}\Delta X$, $\Pi_2 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_{b0}\boldsymbol{\eta}_{-1}$, $\Pi_3 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_{b0}\mathbf{W}_1\boldsymbol{\eta}$, $\Pi_4 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_{b0}\mathbf{W}_2\boldsymbol{\eta}_{-1}$,
$\Phi_1 = \frac{1}{2\sigma_{v0}^4}\mathbf{C}^{-1}$, $\Phi_2 = \frac{1}{\sigma_{v0}^2}(\mathbb{C}_{b0}\mathbb{S}_{-1} + \mathbf{B}_{30}^{-1\prime}\mathbf{E}_{\rho 0}\mathbf{B}_{30}^{-1})$, $\Phi_3 = \frac{1}{\sigma_{v0}^2}(\mathbb{C}_{b0}\mathbf{W}_1\mathbb{S} + \mathbf{B}_{30}^{-1\prime}\mathbf{E}_{\lambda_{10}}\mathbf{B}_{30}^{-1})$,
$\Phi_4 = \frac{1}{\sigma_{v0}^2}(\mathbb{C}_{b0}\mathbf{W}_2\mathbb{S}_{-1} + \mathbf{B}_{30}^{-1\prime}\mathbf{E}_{\lambda_{20}}\mathbf{B}_{30}^{-1})$, $\Phi_5 = \frac{1}{\sigma_{v0}^2}[C^{-1} \otimes \big(B_{30}^{-1\prime}(\mathcal{A}_0 - \mathbf{E}_{\lambda_{30}})B_{30}^{-1}\big)]$,
$\Psi_1 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_{b0}\mathbb{R}_{-1}$, $\Psi_2 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_{b0}\mathbf{W}_1\mathbb{R}$, and $\Psi_3 = \frac{1}{\sigma_{v0}^2}\mathbb{C}_{b0}\mathbf{W}_2\mathbb{R}_{-1}$.

The similarity between (2.4) and (3.3) immediately leads to an MD representation for the CH-robust AQS function, i.e., $S_{\mathtt{H}}^{*}(\psi_0) = \sum_{i=1}^{n}\mathbf{g}_{\mathtt{H}i}$, referring to (2.5)-(2.10). The vectors $S_{\mathtt{H}}^{*}(\psi)$ and $\mathbf{g}_{\mathtt{H}i}$, and the matrix $\Sigma_{\mathtt{H}}^{*}(\psi) = -\mathrm{E}[\frac{\partial}{\partial\psi'}S_{\mathtt{H}}^{*}(\psi_0)]$ are partitioned in the same way according to ϑ and φ. A similar asymptotic MD representation, as in (2.13)and (2.22), holds for $S_{\mathtt{H},\varphi}^{*\prime}(\tilde{\vartheta}_{\mathtt{H}}, 0_k)$, where $\tilde{\vartheta}_{\mathtt{H}}$ is the constrained estimator under the null. An AQS-based and CH-robust test for testing the hypothesis $H_0 : \varphi = 0$ is thus,

$$T_{\mathtt{M}}^{\dagger} = S_{\mathtt{H},\varphi}^{*\prime}(\tilde{\vartheta}_{\mathtt{H}}, 0_k)\big\{\textstyle\sum_{i=1}^{n}(\tilde{\mathbf{g}}_{\mathtt{H}i,\varphi} - \tilde{\Lambda}_{\mathtt{H}}^{*}\tilde{\mathbf{g}}_{\mathtt{H}i,\vartheta})(\tilde{\mathbf{g}}_{\mathtt{H}i,\varphi} - \tilde{\Lambda}_{\mathtt{H}}^{*}\tilde{\mathbf{g}}_{\mathtt{H}i,\vartheta})'\big\}^{-1}S_{\mathtt{H},\varphi}^{*}(\tilde{\vartheta}_{\mathtt{H}}, 0_k), \tag{3.4}$$

where $\mathtt{M} = \mathtt{PD}$, $\mathtt{DPD}$, $\mathtt{SDPD1}$, $\cdots$, $\mathtt{SDPD5}$, and $\mathtt{SPD}$, associated with the null hypotheses defined in Sec. 1, $\tilde{\Lambda}_{\mathtt{H}}^{*} = \tilde{\Sigma}_{\mathtt{H},\varphi\vartheta}^{*}\tilde{\Sigma}_{\mathtt{H},\vartheta\vartheta}^{*-1}$, and $\tilde{\mathbf{g}}_{\mathtt{H}i,\vartheta}$ and $\tilde{\mathbf{g}}_{\mathtt{H}i,\varphi}$ are the null estimates of $\mathbf{g}_{\mathtt{H}i,\vartheta}$ and $\mathbf{g}_{\mathtt{H}i,\varphi}$. We take $\tilde{\Sigma}_{\mathtt{H}}^{*} = -\frac{\partial}{\partial\psi}S_{\mathtt{H}}^{*}(\psi)|_{\psi=\tilde{\psi}_{\mathtt{H}}}$ with $\frac{\partial}{\partial\psi}S_{\mathtt{H}}^{*}(\psi)$ being given in Appendix B.

Theorem 3.1. *Under Assumptions A, B*, C and D, if $\tilde{\vartheta}_{\mathtt{H}}$ is $\sqrt{N}$-consistent, we have under $H_0^{\mathtt{M}}$, $T_{\mathtt{M}}^{\dagger} \xrightarrow{D} \chi_k^2$, as $n \to \infty$, where $\mathtt{M}$ denotes a null model specified in Sec. 1.*

Working with the numerator of $S_{\mathtt{H}}^{*c}(\delta)$ given in (3.2), one may be able to obtain finite sample improved tests that are fully robust against unknown CH. However, this does not seem to be an easy task, as the existence of unknown CH renders the simple recentering method followed in Sec. 2.3 for the homoskedastic case unapplicable. This is seen from the results given in Li and Yang (2020b): $\mathrm{E}(\Delta Y_{-1}\Delta v') = -\sigma_{v0}^2\mathbf{D}_{-10}\mathbf{B}_{30}^{-1}\mathbf{H}$ and $\mathrm{E}(\Delta Y\Delta v') = -\sigma_{v0}^2\mathbf{D}_0\mathbf{B}_{30}^{-1}\mathbf{H}$, where $\mathbf{H} = I_{T-1}\otimes\mathcal{H}_n$ and $\mathcal{H}_n = \mathrm{diag}\{h_{ni}, i = 1, \ldots, n\}$.

4. Monte Carlo Simulation

Monte Carlo experiments are carried out to investigate the finite sample performance of the proposed AQS test $T_{\mathtt{M}}$, standardized AQS (SAQS) test $T_{\mathtt{M}}^{\diamond}$, and the CH-robust AQS test $T_{\mathtt{M}}^{\dagger}$, in terms of size and size-adjusted power of the tests. The following data generating process (DGP) is followed:

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

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 (2018a), and the elements of Z are randomly generated from $Bernoulli(0.5)$.

The spatial weight matrices are generated according to `Rook` contiguity, `Queen` contiguity, or group interaction schemes: `Group-I` or `Group-II`.[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 (for $T = 3$) or 2000 (for $T = 6$). The error (v_{it}) distributions can be (i) normal, (ii) normal mixture (10% $N(0, 4)$ and 90% $N(0, 1)$), or (iii) lognormal.[7] The fixed effects μ are generated according to $\frac{1}{T}\sum_{t=1}^{T} X_t + e$, where $e \sim (0, I_N)$. The cross-sectional heteroskedasticity (CH) is generated according to: `CH-1`: $h_{ni} \propto \frac{1}{T}\sum_{t=1}^{T} |\Delta X_{nt}|$; `CH-2`: $h_{ni} \propto n_g$ for ith unit in gth group of size n_g; and `CH-3`: $h_{ni} \propto n_g$ if $n_g \leq \bar{n}_k$; and $\propto 1/n_g^2$ otherwise, where $\bar{n}_k$ is the average group size. The case of homoskedasticity is denoted as `CH-0`. `Group-I` gives strongest spatial interaction and `CH-3` gives the most severe cross-sectional heteroskedasticity. Under `Group-II`, variation in number of neighbors for each spatial unit stays constant as n increases; in all other spatial layouts, it vanishes as n increases although slower for `Group-I` (see Yang, 2010).

We report in Tables 1a-1c partial results for testing $H_0^{\texttt{PD}}$: $\delta = 0$. When n is not large, the AQS test $T_{\texttt{PD}}$ and the CH-robust AQS (RAQS) test $T_{\texttt{PD}}^{\dagger}$ can be severely oversized, whereas the standardized AQS (SAQS) test $T_{\texttt{PD}}^{\diamond}$ can be slightly undersized. As n increases, the empirical sizes of $T_{\texttt{PD}}^{\diamond}$ quickly approach to their nominal values corresponding to the χ_4^2 distribution. As T increases from 3 (Table 1a) to 6 (Table 1b), all tests improve significantly. As shown by Corollary 2.1 and Theorem 3.1, these tests are all robust against unknown CH. The results given in Table 1b confirm this. The results further reveal that the severity of CH has a much greater impact on the AQS and RAQS tests than on the SAQS test in finite sample performance. As all three tests are asymptotically valid, it is important to compare their finite sample performance in terms of the power of the tests. This has to be done with sizes being adjusted. The results in Table 1c show that the size-adjusted power is the highest for $T_{\texttt{PD}}^{\diamond}$ and the lowest for $T_{\texttt{PD}}^{\dagger}$, as expected.

Tables 2a-2c present partial results for testing $H_0^{\texttt{DPD}}$: $\lambda = 0$, allowing ρ to be present in the model as a free parameter. The results show an excellent performance of the SAQS test with its empirical sizes being very close to their nominal values even when $n = 50$. In contrast, the regular and robust AQS tests can have sever size distortions when n is not so large, which get smaller in a significantly slower speed than those of the SAQS test, in particular under CH. While all three tests are robust against unknown CH as shown by Corollary 2.2 and Theorem 3.1, their finite sample properties differ (from both reported and unreported results), with $T_{\texttt{DPD}}$ and $T_{\texttt{DPD}}^{\dagger}$ being affected by the severity of CH much more than the SAQS test $T_{\texttt{DPD}}^{\diamond}$. When T increases from 3 to 6, the AQS and RAQS tests improve significantly. The SAQS test is in general slightly more powerful than the AQS and RAQS tests. The true value of ρ does not have a significant effect on both tests.

We now turn to the tests of $H_0^{\texttt{SDPD4}}$: $\lambda_1 = \lambda_2 = 0$, allowing ρ and λ_3 to be present in the model as free parameters. As this is a case under which the null model contains both spatial and dynamic parameters and the AQS and SAQS tests are non-robust against unknown CH, we therefore focus on the two main issues: the performance of the tests when the spatial parameters approach to the boundary of parameter space,[8] and the severity of CH that would lead to RAQS to perform better than AQS and SAQS.

[6] The `Rook` and `Queen` schemes are standard. For `Group-I`, 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 $\sum_{g=1}^{k} n_g = n$. For `Group-II`, we first generate 6 groups of fixed sizes $(3, 5, 7, 9, 11, 15)$, and replicate these groups r times to give $n = r \times 50$. See Lin and Lee (2010) and Yang (2018a) for details in generating these spatial layouts.

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

[8] See Lee and Yu (2016) for a detailed discussion on parameter identification of the SDPD model.

Table 3a presents results with $\lambda_3 = \pm.9$, close to the upper or lower boundaries of the parameter space for λ_3 with W_3 being row-normalized. The results show that the value of λ_3 does not affect much on the performance of the tests. The results (reported and unreported) further show that under the homoskedasticity, these tests perform reasonably well, although not as well as the SAQS tests for the first two cases and some unreported cases. Table 3b presents partial results based on `Group-I` and `Group-II` spatial layouts, and under `CH-1`, `CH-2` and `CH-3`. The results show that the empirical sizes of the RAQS test generally converge to their nominal levels no matter how severe the CH is, showing the full robustness of the RAQS test. For AQS and SAQS tests, the results show that they can be quite robust against mild CH, but under the most severe CH (`CH-3` with `Group-II`) their size distortions do not get smaller when sample size becomes larger. In contrast, the size distortions of the RAQS test almost vanishes at $n = 500$.

Additional Monte Carlo results for the three reported cases and several unreported cases are given in a `Supplementary Appendix` to this paper available at author's website: `http://www.mysmu.edu/faculty/zlyang/`. All results suggest that if a test has a PD or a DPD as its null model, the SAQS test is recommended as it has a much better finite sample performance than the other two, given that all three tests are robust against unknown CH. The results also show that in many situations the AQS and SAQS tests are quite robust against mild departure from homoskedasticity of the errors, and the SAQS tests compare favorably against the RAQS tests. In a situation under which the AQS and SAQS are more sensitive to CH and/or when heteroskedasticity is truly in doubt, the fully robust version of the tests may be used. Finally, when the null model involves spatial and/or dynamic parameters with their true values being close to the boundary of the parameter space, how do the corresponding tests perform? Our results suggest that λ_3 may have a bigger impact on the performance of the tests than the other three.

5. 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 adjusted quasi scores (AQS) and their martingale difference representations. Standardized versions of the AQS tests are also introduced, by adjusting the concentrated quasi scores, for an improved finite sample performance. The standardized versions of the tests are shown to be as simple as the non-standardized versions but are more reliable in finite samples and are quite robust against the unknown CH in general, hence are recommended for the empirical applications. In case of severe cross-sectional heteroskedasticity (CH) and when the regular AQS tests are non-robust, the AQS tests fully robust against unknown CH are also introduced. Monte Carlo results show excellent performance of the standardized AQS tests and the full robustness of the robust versions of the AQS tests under severe CH. The results presented in the paper show that the general methodology for constructing tests of this nature are promising – it overcomes the difficulty faced by the short (spatial) dynamic panel models.

Appendix A: Some Useful Lemmas

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 $\{A_nB_n\}$ 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_nC_n and C_nA_n are uniformly $O(\iota_n^{-1})$.

Lemma A.2. (Lee, 2004, p.1918). *For W_r and B_r, $r = 1, 3$, defined in Model* (1.1), *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 λ_{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. (Li and Yang, 2020b) *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 $O(\iota^{-1})$ uniformly in all i and j. Let v_n be a random n-vector of inid elements satisfying Assumption B, and b_n a constant n-vector of elements of uniform order $O(\iota^{-1/2})$. Then*

(i) $\mathrm{E}(v_n'A_nv_n) = O(\frac{n}{\iota_n})$, *(ii)* $\mathrm{Var}(v_n'A_nv_n) = O(\frac{n}{\iota_n})$,
(iii) $\mathrm{Var}(v_n'A_nv_n + b_n'v_n) = O(\frac{n}{\iota_n})$, *(iv)* $v_n'A_nv_n = O_p(\frac{n}{\iota_n})$,
(v) $v_n'A_nv_n - \mathrm{E}(v_n'A_nv_n) = O_p((\frac{n}{\iota_n})^{\frac{1}{2}})$, *(vi)* $v_n'A_nb_n = O_p((\frac{n}{\iota_n})^{\frac{1}{2}})$,

and (vii), the results (*iii*) *and* (*vi*) *remain valid if b_n is a random n-vector independent of v_n such that $\{\mathrm{E}(b_{ni}^2)\}$ are of uniform order $O(\iota_n^{-1})$.*

Lemma A.5. (Li and Yang, 2020b)*: Let $\{\Phi_n\}$ be a sequence of $n \times n$ matrices with row and column sums uniformly bounded, and elements of uniform order $O(\iota_n^{-1})$. Let v_n be a random n-vector satisfying Assumption B. Let $b_n = \{b_{ni}\}$ be an $n \times 1$ random vector, independent of v_n, such that (i) $\{\mathrm{E}(b_{ni}^2)\}$ are of uniform order $O(\iota_n^{-1})$, (ii) $sup_i E|b_{ni}|^{2+\epsilon_0} < \infty$, (iii) $\frac{\iota_n}{n}\sum_{i=1}^n[\phi_{n,ii}(b_{ni} - \mathrm{E}b_{ni})] = o_p(1)$ where $\{\phi_{n,ii}\}$ are the diagonal elements of Φ_n, and (iv) $\frac{\iota_n}{n}\sum_{i=1}^n[b_{ni}^2 - \mathrm{E}(b_{ni}^2)] = o_p(1)$. Let $\mathcal{H}_n = \mathrm{diag}(h_{n1}, \ldots, h_{nn})$. Define the bilinear-quadratic form:*

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

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 $Q_n/\sigma_{Q_n} \xrightarrow{d} N(0, 1)$.

The following lemma extends the formulations in Sec. 2.2 to allow for unknown CH. Its proof follows the results of Theorems 3.2 and 3.3 of Li and Yang (2020b). Recall: A^u, A^l and A^d denote the upper-triangular, lower-triangular, and diagonal matrix of a square matrix A; Π_t, Φ_{ts} and Ψ_{ts} the submatrices of Π, Φ and Ψ partitioned according to $t, s = 2, \ldots, T$; $\Psi_{t+} = \sum_{s=2}^T \Psi_{ts}$, $\Theta = \Psi_{2+}(B_{30}B_{10})^{-1}$, $\Delta y_1^\circ = B_{30}B_{10}\Delta y_1$, and $\Delta y_{1t}^* = \Psi_{t+}\Delta y_1$.

Lemma A.6. *Suppose Assumptions A, B*, C-E hold for Model* (2.1). *Consider the linear, quadratic and bilinear forms, $Q(\psi_0) = \{(\Pi\Delta'v)', \Delta v'\Phi\Delta v, (\Delta v'\Psi\Delta\mathbf{y}_1)'\}'$, associated with the model. Assume the elements of Π ($N \times 1$) are uniformly bounded, and the matrices Φ and Ψ ($N \times N$) are uniformly bounded in both row and column sums. 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* $\mathrm{diag}\{\Theta_{ii}\} = \Theta^d$. *Then, 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*) $\frac{1}{N}[Q(\psi_0) - \mathrm{E}(Q(\psi_0))] \xrightarrow{D} N(0, \lim_{n\to\infty} \frac{1}{N}\Gamma)$, *where* $\Gamma = \mathrm{Var}(Q(\psi_0))$.

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

Appendix B: Some Technical Details

We sketch the proofs of the theorems, corollaries, and lemmas. Details are given in `Supplementary Appendix` available at `http://www.mysmu.edu/faculty/zlyang/`, including the detailed expressions for the derivative matrices of the three AQS functions, which is referred to loosely as the Hessian matrix in this paper.

Proof Theorem 2.1: The proof of Theorem 2.1 follows closely the proofs of Theorems 3.2 and 3.3 of Yang (2018a), and is typically simpler as under $H_0^{\mathtt{M}}$ the model becomes simpler. The Hessian matrix $\frac{\partial}{\partial\psi'}S^*(\psi)$ used to estimate $\Sigma^*(\psi_0)$ can be easily derived based on the expression of $S^*(\psi)$ given in (2.2). It can found in Yang (2018a, Proof of Theorem 3.2), and also in the `Supplementary Appendix` to this paper containing additional 'asymmetric components' that did not appear in Yang (2018a).

Proof of Corollary 2.1. The quantities needed for evaluating the AQS function defined in (2.4) become: $\Pi_1 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\Delta X$, $\Pi_2 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbb{B}_{-1}\Delta X\beta$, $\Pi_3 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_1\Delta X\beta$, $\Pi_4 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_2\mathbb{B}_{-1}\Delta X\beta$, $\Phi_1 = \frac{1}{2\sigma_{v0}^4}\mathbf{C}^{-1}$, $\Phi_2 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbb{B}_{-1}$, $\Phi_3 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_1$, $\Phi_4 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_2\mathbb{B}_{-1}$, $\Phi_5 = \frac{1}{2\sigma_{v0}^2}[C^{-1}\otimes(W_3'+W_3)]$, $\Psi_1 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbb{R}_{-1}$, $\Psi_2 = 0$, $\Psi_3 = \frac{1}{\sigma_{v0}^2}\mathbf{C}^{-1}\mathbf{W}_2\mathbb{R}_{-1}$, $\mathbb{R}_{-1} = \mathrm{blkdiag}(I_n, 0, \ldots, 0)$, $\mathbb{B}_{-1} = I_{T-1}^* \otimes I_n$, 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 Φ and Ψ matrices are either of the form $A\otimes I_n$ or $A \otimes W$ for some $(T-1)\times(T-1)$ matrix A and a spatial weight matrix W satisfying Assumption D. Thus, $\mathrm{E}[S^*(\psi_0)] = 0$ even when the errors are heteroskedastic. Hence by Lemma A.5, we have $\frac{1}{\sqrt{N}}S^*(\psi_0) \xrightarrow{D} N[0, \lim_{n\to\infty}\frac{1}{N}\Gamma^*(\psi_0)]$.

By the mean value theorem (MVT), 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 (2.12) $S^*(\psi_0) = \sum_{i=1}^n \mathbf{g}_i$, where $\{\mathbf{g}_i, \mathcal{F}_{n,i}\}$ form a vector MD sequence, we have $\frac{1}{N}\sum_{i=1}^n[\mathbf{g}_i\mathbf{g}_i' - \mathrm{E}(\mathbf{g}_i\mathbf{g}_i')] = o_p(1)$ by Lemma A.6. By MVT 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. Finally, it is easy to show that $\mathrm{plim}_{n\to\infty}(\widetilde\Lambda - \Lambda) = 0$, using the simplified expression of $H^*(\psi)$ and MVT.

Proof of Lemma 2.1. Consider the AQS vector $S^*(\beta, \sigma_v^2, \rho)$ for the DPD model, and the concentrated AQS function which defines $\tilde\rho$ under $H_0^{\mathtt{DPD}}$:

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

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

Define $\bar{S}^*(\beta, \sigma_v^2, \rho) = \mathrm{E}[S^*(\beta, \sigma_v^2, \rho)]$. Given ρ, $\bar{S}^*(\beta, \sigma_v^2, \rho) = 0$ is partially solved at $\bar{\beta}(\rho) = (\Delta X'\mathbf{C}^{-1}\Delta X)^{-1}\Delta X'\mathbf{C}^{-1}(\mathrm{E}\Delta Y - \rho\mathrm{E}\Delta Y_{-1})$ and $\bar{\sigma}_v^2(\rho) = \frac{1}{N}\mathrm{E}[\Delta\bar{v}(\rho)'\mathbf{C}^{-1}\Delta\bar{v}(\rho)]$, 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 $S^{*c}_{\mathtt{DPD}}(\rho)$ as

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

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.

Proof of Corollary 2.2. First, with $\psi_0 = (\beta_0, \sigma_{v0}^2, \rho_0, 0_3')'$ it is easy to show that $\mathrm{E}[S^*(\psi_0)] = 0$ under the general heteroskedasticity $\{h_{ni}\}$. By Lemma A.5, one shows that $\frac{1}{\sqrt{N}}S^*(\psi_0) \xrightarrow{D} N(0, \Gamma^*(\psi_0))$. By Lemma A.6, 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 $\tilde{\beta}$, $\tilde{\sigma}_v^2$ and $\tilde{\rho}$ against unknown heteroskedasticity $\{h_{ni}\}$ as shown in Lemma 2.1, we have $\frac{1}{\sqrt{N}}[S^*_\lambda(\tilde{\psi}) - S^*_\lambda(\psi_0)] \xrightarrow{p} 0$ where $\tilde{\psi} = (\tilde{\beta}', \tilde{\sigma}_v^2, \tilde{\rho}, 0_3)'$, and $\frac{1}{N}\sum_{i=1}^n(\tilde{\mathbf{g}}_{n,i}\tilde{\mathbf{g}}_{n,i}' - \mathbf{g}_i\mathbf{g}_i') \xrightarrow{p} 0$. Finally, using the simplified expression of $H^*(\psi)$ and MVT, we show $\mathrm{plim}_{n\to\infty}(\widetilde{\Lambda} - \Lambda) = 0$.

Proof of Theorem 2.2: The proof is similar to that of Theorem 2.1. The partial derivatives of $S^\diamond(\delta)$ required to estimate the components of $\Sigma^\diamond_{\varphi\pi}(\delta_0)$ and $\Sigma^\diamond_{\varphi\varphi}(\delta_0)$ can be easily obtained from the expression $S^\diamond(\delta)$ given in (2.19). The full expression of $\frac{\partial}{\partial\delta'}S^\diamond(\delta)$ is given in the `Supplementary Appendix` to this paper.

Proof of Theorem 3.1: The proof of Theorem 3.1 follows closely the proofs of Theorems (3.2) and (3.3) of Li and Yang (2020b). The Hessian matrix $\frac{\partial}{\partial\psi'}S^*_{\mathtt{H}}(\psi)$ used to estimate $\Sigma^*_{\mathtt{H}}(\psi_0)$ is given in Li and Yang (2020b, Proof of Theorem 3.2), and can also be found in the `Supplementary Appendix` to this paper, where the 'asymmetric components' that did not appear in Li and Yang (2020b) are also given.

Compliance with Ethical Standards:

`Conflict of Interest:` The author declares that he has no conflict of interest.
`Ethical approval:` This article does not contain any studies with human participants or animals performed by the author.

References

[1] Anselin, L., Bera, A. K., Florax, R., Yoon, M. J., 1996. Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics* 26, 77-104.

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

[3] 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.

[4] Bai, J., Li, K., 2015. Dynamic spatial panel data models with common shocks. *Working paper*, Columbia University, New York.

[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] Bera, A. K., Dogan, O., Taspinar, S., Leiluo Y., 2019. Robust LM tests for spatial dynamic panel data models. *Regional Science and Urban Economics* 76, 47-66.

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

[11] Chudik, A., Pesaran, M. H., 2017. A bias-corrected method of moments approach to estimation of dynamic short-T panels. *CESifo Working Paper Series* No. 6688.

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

[13] Elhorst, J. P., 2010. Dynamic panels with endogenous interaction effects when T is small. *Regional Science and Urban Economics* 40, 272-282.

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

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

[16] Jin, F., Lee, L. F., 2018. Outer-product-of-gradients tests for spatial autoregressive models. *Regional Science and Urban Economics* 72, 35-57.

[17] 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.

[18] 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.

[19] Kuersteiner, G. M., Prucha, I. R., 2018. Dynamic panel data models: networks, common shocks, and sequential exogeneity. *Working Paper*, University of Maryland, College Park.

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

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

[22] Lee, L. F., Yu, J., 2010. A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory* 26, 564-597.

[23] Lee, L. F., Yu, J., 2012. Spatial panels: random components vs. fixed effects. *International Economic Review* 53, 1369-1412.

[24] Lee, L. F., Yu, J., 2014. Efficient GMM estimation of spatial dynamic panel data models with fixed effects. *Journal of Econometrics* 180, 174-197.

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

[26] Li, L. Y., Yang, Z. L. 2020a. Spatial dynamic panel data models with correlated random effects. *Journal of Econometrics*, forthcoming.

[27] Li, L. Y., Yang, Z. L. 2020b. Estimation of fixed effects spatial dynamic panel data models with small T and unknown heteroskedasticity. *Regional Science and Urban Economics*, forthcoming.

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

[29] 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.

[30] Liu, X., Prucha, I. R., 2018. A robust test for network generated dependence. *Journal of Econometrics* 207, 92-113.

[31] Mutl, J., 2006. Dynamic panel data models with spatially correlated disturbances. *PhD Thesis*, University of Maryland, College Park.

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

[33] Qu, X., Wang, X., Lee, L. F., 2016. Instrumental variable estimation of a spatial dynamic panel model with endogenous spatial weights when T is small. *Econometrics Journal* 19, 261-290.

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

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

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

[37] Shi, W., Lee, L. F., 2017. Spatial Dynamic Panel Data Models with Interactive Fixed Effects. *Journal of Econometrics* 197, 323-347.

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

[39] Taspinar, S., Dogan, O., Bera, A. K., 2017. GMM gradient tests for spatial dynamic panel data models. *Regional Science and Urban Economics* 65, 65-88.

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

[41] Xu, Y. H., Yang, Z. L. (2019). Specification tests for temporal heterogeneity in spatial panel data models with fixed effects. *Regional Science and Urban Economics*, forthcoming.

[42] Yang, Z. L., Li, C., Tse, Y. K., 2006. Functional form and spatial dependence in dynamic panels. *Economics Letters* 91, 138-145.

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

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

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

[46] Yang, Z. L., 2018b. `Supplement` to "Unified M-estimation of fixed-effects spatial dynamic models with short panels. *Journal of Econometrics* 205, 423-447", `http://www.mysmu.edu/faculty/zlyang/`

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

[48] Yu, J., de Jong, R., Lee, L. F., 2008. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics* 146, 118-134.

[49] Yu, J., Lee, L. F., 2010. Estimation of unit root spatial dynamic panel data models. *Econometric Theory* 26, 1332-1362.

Table 1a Empirical Size of Tests of $H_0^{\texttt{PD}} : \delta = 0$; `Group-I`, $T = 3$

			$T_{\texttt{PD}}$			$T_{\texttt{PD}}^{\diamond}$			$T_{\texttt{PD}}^{\dagger}$		
`CH`	n	dgp	10%	5%	1%	10%	5%	1%	10%	5%	1%
`CH-0`	50	1	15.30	8.08	1.22	10.78	4.28	0.46	15.54	8.20	1.16
		2	22.80	14.18	3.90	8.46	3.20	0.22	25.76	16.60	5.18
		3	16.48	9.46	2.16	8.78	3.44	0.22	18.24	10.54	2.64
	100	1	12.60	6.48	1.28	10.42	5.08	0.60	13.30	7.38	1.56
		2	17.22	10.34	3.16	9.78	4.16	0.72	19.68	12.16	3.94
		3	14.00	7.96	2.10	9.80	4.54	0.80	16.04	8.98	2.50
	200	1	11.14	6.52	1.26	10.56	5.30	0.90	12.66	6.84	1.50
		2	14.08	7.74	1.90	9.22	4.18	0.58	16.34	9.18	2.80
		3	13.70	7.20	1.76	10.78	5.12	0.78	14.60	8.14	1.96
	500	1	10.78	5.58	1.32	10.34	4.94	1.28	11.80	6.22	1.46
		2	12.66	6.96	1.44	10.38	5.16	0.80	13.30	7.16	2.00
		3	11.98	6.40	1.62	10.82	5.50	1.08	13.22	7.14	1.56
`CH-1`	50	1	19.40	11.52	2.96	9.86	3.82	0.30	19.44	11.92	2.90
		2	27.20	18.52	6.98	8.18	2.96	0.16	29.78	21.16	9.34
		3	22.42	13.94	4.64	8.26	3.28	0.22	24.26	15.10	4.94
	100	1	15.70	9.12	2.46	10.12	4.52	0.52	15.46	9.42	2.30
		2	22.14	14.34	4.98	8.86	3.74	0.32	25.44	16.30	6.18
		3	18.52	11.10	3.50	9.66	4.28	0.64	20.50	13.00	4.30
	200	1	14.00	7.54	1.72	11.08	5.34	0.76	14.10	7.78	2.14
		2	17.08	9.88	2.64	9.50	3.94	0.56	19.02	11.32	3.74
		3	14.72	8.24	2.10	9.86	4.56	0.90	15.78	8.94	2.38
	500	1	11.44	5.94	1.42	10.54	5.24	1.16	12.30	6.52	1.50
		2	12.84	6.98	1.44	9.12	4.18	0.60	14.84	8.66	2.42
		3	11.32	6.04	1.20	9.74	4.70	0.74	13.48	7.20	1.70
`CH-2`	50	1	15.08	8.42	1.74	10.74	4.72	0.58	16.08	9.02	2.18
		2	21.58	13.04	4.10	8.58	3.40	0.20	23.90	15.06	5.36
		3	17.44	9.94	2.64	9.16	3.58	0.40	19.04	10.98	2.98
	100	1	12.26	6.68	1.58	10.42	4.86	0.90	13.72	7.04	1.82
		2	17.52	9.96	2.98	9.52	4.18	0.52	20.08	12.68	4.06
		3	14.42	7.94	2.18	10.32	4.42	0.62	15.76	9.22	2.26
	200	1	11.36	5.94	1.18	10.08	4.86	0.74	12.44	6.72	1.24
		2	14.48	8.98	2.26	10.12	4.78	0.74	15.78	9.48	2.58
		3	13.74	7.80	1.92	11.20	5.78	0.78	15.52	8.72	2.00
	500	1	10.74	5.56	1.02	10.26	5.06	0.86	12.58	6.68	1.30
		2	11.34	5.82	1.40	9.60	4.66	0.86	13.52	7.66	1.94
		3	11.04	5.84	1.56	10.04	4.96	1.24	12.68	7.02	1.74
`CH-3`	50	1	23.94	14.86	4.80	7.74	2.64	0.18	26.34	16.88	5.80
		2	33.68	24.62	12.16	6.26	1.94	0.00	39.92	30.00	15.90
		3	28.44	19.60	8.26	7.06	2.18	0.02	32.94	23.00	10.14
	100	1	22.80	14.76	5.32	9.64	3.78	0.10	26.88	18.24	7.14
		2	31.50	22.62	11.42	7.40	2.72	0.08	40.06	30.60	16.96
		3	26.26	18.16	7.78	7.90	2.84	0.20	33.40	24.02	11.40
	200	1	15.44	9.04	2.72	10.42	4.56	0.72	17.18	10.80	3.50
		2	22.70	14.42	5.44	9.48	3.64	0.28	26.18	18.00	7.62
		3	19.08	11.78	3.88	10.28	4.82	0.54	21.08	13.38	4.70
	500	1	13.48	7.48	1.92	10.96	5.26	0.94	14.90	8.34	2.34
		2	16.76	9.84	2.96	9.10	4.02	0.58	19.86	12.16	4.52
		3	14.32	8.20	2.26	10.14	4.52	0.86	17.74	11.04	3.44

Note: for dpg, 1 = normal, 2 = normal mixture, 3 = lognormal

Table 1b Empirical Size of Tests of $H_0^{\text{PD}}: \delta = 0$; `Group-I`, $T = 6$

			T_{PD}			$T_{\text{PD}}^{\diamond}$			$T_{\text{PD}}^{\dagger}$		
`CH`	n	dgp	10%	5%	1%	10%	5%	1%	10%	5%	1%
`CH-0`	50	1	11.00	5.45	0.70	9.00	4.05	0.45	10.00	5.10	0.90
		2	17.05	10.20	2.20	9.00	3.50	0.35	17.60	11.05	2.55
		3	13.45	6.80	1.60	9.25	4.35	0.60	14.45	7.30	1.85
	100	1	11.20	5.75	1.25	10.15	5.25	0.90	11.65	5.60	1.20
		2	13.55	7.65	1.55	9.75	4.35	0.55	12.65	6.55	1.90
		3	12.25	6.15	1.30	9.60	4.50	0.70	12.75	6.50	1.25
	200	1	10.40	4.95	0.65	9.85	4.55	0.60	10.40	4.55	0.70
		2	13.10	7.30	1.60	11.00	5.00	0.95	13.00	7.50	1.75
		3	10.65	4.95	0.95	9.30	4.15	0.50	11.65	4.90	1.10
	500	1	10.25	5.15	0.85	10.10	4.90	0.80	11.05	5.85	1.05
		2	11.25	6.30	1.30	10.35	5.50	1.00	11.65	6.35	1.15
		3	9.85	4.70	0.95	9.20	4.30	0.70	10.55	5.50	1.00
`CH-1`	50	1	13.40	6.90	1.90	8.35	4.50	0.25	13.65	7.00	1.80
		2	19.25	12.30	3.10	8.25	3.05	0.10	21.00	13.25	4.00
		3	15.95	9.20	1.75	8.55	3.55	0.05	17.45	10.10	2.90
	100	1	10.20	5.60	1.05	8.20	3.85	0.60	12.30	6.95	1.25
		2	15.70	9.05	2.45	9.70	4.25	0.50	16.95	10.75	3.25
		3	13.90	8.25	2.20	9.55	4.55	0.65	15.10	8.20	2.00
	200	1	11.85	5.80	1.00	10.20	4.65	0.75	11.50	6.00	1.15
		2	13.00	7.00	1.70	9.15	4.15	0.70	14.30	7.80	2.25
		3	11.75	6.30	1.45	9.75	4.20	1.00	12.65	6.65	1.85
	500	1	10.65	4.90	1.00	9.65	4.75	0.75	10.95	5.60	1.25
		2	12.05	6.45	1.55	10.05	5.10	1.00	13.60	7.60	1.55
		3	10.60	5.50	1.05	9.60	4.60	0.70	11.25	6.30	1.95
`CH-2`	50	1	10.10	4.90	0.70	9.30	4.15	0.55	11.45	5.60	1.00
		2	14.65	7.75	1.35	9.15	3.55	0.50	17.30	9.70	2.55
		3	13.25	6.60	0.95	10.10	4.05	0.30	13.75	6.60	1.35
	100	1	11.40	5.50	0.90	10.60	4.85	0.75	11.25	4.90	1.25
		2	14.00	7.10	1.75	9.05	4.35	0.65	13.00	7.75	2.10
		3	12.05	6.35	1.20	10.05	5.00	0.80	13.35	7.20	1.30
	200	1	10.85	5.05	0.75	10.10	4.65	0.55	12.20	5.80	0.90
		2	12.60	6.35	1.85	9.75	4.55	1.20	13.90	7.30	1.90
		3	11.50	5.85	1.25	10.85	4.95	0.70	11.75	6.00	1.60
	500	1	11.15	6.20	1.40	11.00	5.85	1.30	12.15	6.05	1.10
		2	10.85	5.20	1.15	9.70	4.65	0.80	11.25	6.30	1.15
		3	10.35	5.65	1.60	9.75	5.25	1.40	11.45	5.95	1.30
`CH-3`	50	1	19.10	10.65	3.05	8.60	3.45	0.10	18.30	10.60	2.85
		2	26.45	18.00	7.10	7.35	2.95	0.10	31.80	21.15	7.65
		3	23.15	15.30	4.75	8.30	2.70	0.15	23.70	15.20	5.15
	100	1	16.45	9.70	2.35	9.65	3.90	0.55	16.70	10.20	3.50
		2	24.40	15.25	5.75	8.90	3.75	0.35	25.90	17.85	7.45
		3	19.35	11.05	3.90	9.15	3.30	0.10	20.70	12.80	4.30
	200	1	12.40	6.30	1.35	9.15	4.10	0.60	13.25	7.35	1.85
		2	15.70	9.35	3.20	9.30	4.00	0.65	17.50	10.55	3.20
		3	13.80	7.80	1.70	9.35	4.00	0.55	15.05	8.00	1.95
	500	1	12.15	6.55	1.60	10.50	5.55	1.20	12.90	7.20	2.00
		2	13.00	7.00	1.95	9.40	4.85	0.75	13.85	7.40	2.00
		3	10.75	5.85	1.45	9.20	4.20	0.75	12.25	7.05	1.60

Note: for dpg, 1 = normal, 2 = normal mixture, 3 = lognormal

Table 1c Size-Adjusted Power of Tests of $H_0^{\text{PD}}: \delta = 0$; `Group-I`

		T_{PD}			$T_{\text{PD}}^{\diamond}$			$T_{\text{PD}}^{\dagger}$		
n	dgp	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 3$, $H_a^{\text{PD}}: \delta = (.06, .06, .06, .06)'$, `CH-0` (1st panel below); `CH-2` (2nd panel below)										
50	1	31.60	22.16	9.04	31.72	21.22	8.12	31.16	20.44	9.70
	2	28.10	18.30	4.58	40.30	29.90	12.20	26.62	16.88	5.00
	3	30.48	19.78	6.66	34.80	23.30	8.26	26.70	17.40	5.48
100	1	44.38	31.48	13.58	43.52	29.92	12.82	39.44	28.46	11.56
	2	39.06	25.02	8.46	51.62	38.44	14.44	33.60	21.14	7.66
	3	43.00	28.96	10.46	48.28	34.12	12.74	37.20	25.78	8.64
200	1	69.44	56.70	31.42	68.96	56.36	31.90	61.18	46.86	24.82
	2	64.40	50.90	26.14	72.16	61.94	40.02	49.26	36.44	16.38
	3	65.36	53.18	26.60	69.20	57.04	34.50	55.24	42.54	21.24
500	1	98.24	96.40	85.24	98.20	96.38	84.60	95.02	90.84	74.78
	2	96.56	92.98	78.96	97.76	95.74	87.26	83.16	74.48	53.64
	3	97.86	95.08	84.28	98.22	96.22	87.18	91.92	86.20	71.82
50	1	33.08	21.38	7.50	31.90	20.42	5.74	31.60	20.24	6.38
	2	29.14	17.90	4.74	40.36	29.32	11.24	26.24	15.84	4.46
	3	30.50	19.30	6.36	34.52	23.20	8.88	26.84	17.32	5.52
100	1	40.48	27.70	9.64	40.44	27.94	10.04	35.76	24.20	7.56
	2	34.92	22.44	8.08	43.94	31.72	13.30	27.62	17.52	6.68
	3	36.72	24.66	8.24	40.06	28.14	11.44	29.82	19.98	7.66
200	1	68.80	56.16	33.44	68.58	55.72	33.00	59.64	47.14	26.54
	2	60.18	46.66	23.88	69.28	57.38	36.46	48.94	37.10	17.30
	3	62.50	50.28	28.74	66.80	54.68	32.78	50.22	38.42	20.86
500	1	96.76	94.24	83.50	96.84	94.28	83.52	91.54	85.50	67.94
	2	94.62	89.36	73.20	96.56	92.92	81.06	74.86	62.48	39.32
	3	96.48	93.00	76.98	97.12	94.76	81.48	88.02	79.56	58.14
$T = 6$, $H_a^{\text{PD}}: \delta = (.05, .05, .05, .05)'$, `CH-0` (1st panel below); `CH-2` (2nd panel below)										
50	1	63.00	49.45	26.85	62.05	47.40	24.70	55.90	42.50	20.40
	2	56.60	41.90	14.25	72.20	56.30	29.80	43.70	30.85	11.60
	3	62.40	48.75	20.20	66.75	55.50	28.55	51.85	37.05	15.15
100	1	80.15	70.05	46.85	79.95	70.15	43.30	64.85	51.65	25.20
	2	74.70	63.30	34.95	83.25	74.70	52.55	53.70	38.55	16.50
	3	76.25	65.85	39.35	78.75	69.25	44.60	54.00	37.60	16.20
200	1	98.00	96.50	85.60	98.10	96.60	84.45	92.85	86.65	67.75
	2	96.95	93.85	81.50	98.40	96.40	89.65	80.95	72.30	45.25
	3	98.15	95.80	85.50	98.90	96.95	89.65	87.25	78.75	55.10
500	1	100.00	100.00	99.90	100.00	100.00	99.90	99.80	99.30	96.70
	2	99.90	99.85	99.45	99.95	99.90	99.80	96.15	92.85	82.20
	3	100.00	99.90	99.90	100.00	99.90	99.90	99.10	97.90	91.90
50	1	50.80	35.45	11.70	50.45	36.50	11.85	37.95	24.00	6.80
	2	43.35	28.85	12.05	56.25	43.90	22.40	27.25	16.60	4.25
	3	46.10	32.90	14.25	51.80	37.90	18.65	31.15	20.25	6.45
100	1	74.60	64.75	40.45	74.55	64.25	40.95	59.85	46.75	25.80
	2	72.25	57.95	30.05	79.05	70.35	45.95	46.05	31.70	14.50
	3	75.55	62.85	38.80	78.70	68.05	43.65	51.50	36.30	15.60
200	1	98.15	96.30	89.75	98.30	96.30	90.00	90.20	85.75	67.75
	2	96.25	93.35	76.35	97.70	96.00	87.50	78.10	65.45	44.05
	3	97.60	95.10	85.60	98.25	96.35	88.45	83.80	75.35	54.85
500	1	100.00	100.00	100.00	100.00	100.00	100.00	99.75	99.55	98.10
	2	99.95	99.90	99.75	100.00	99.95	99.85	96.30	92.50	78.65
	3	99.95	99.80	99.70	100.00	99.85	99.75	99.00	97.60	90.85

Note: for dpg, 1 = normal, 2 = normal mixture, 3 = lognormal

Table 2a Size and Size-Adjusted Power of Tests of $H_0^{\texttt{DPD}} : \lambda = 0$; `Rook`, $\rho = 0.5$, $T = 3$

		$T_{\texttt{DPD}}$			$T^{\diamond}_{\texttt{DPD}}$			$T^{\dagger}_{\texttt{DPD}}$		
n	dgp	10%	5%	1%	10%	5%	1%	10%	5%	1%
Size, `CH-0` (1st panel below); `CH-1` (2nd panel below)										
50	1	12.42	6.20	1.30	9.56	4.42	0.52	12.38	6.40	1.44
	2	16.50	9.90	2.44	8.74	3.56	0.28	16.36	10.00	2.38
	3	15.40	8.50	2.06	9.86	4.10	0.46	15.30	8.48	2.08
100	1	11.12	5.88	1.32	10.02	4.82	0.82	11.32	5.96	1.38
	2	14.36	7.66	1.92	9.10	3.84	0.70	14.32	7.96	2.00
	3	13.00	6.46	1.60	9.34	4.40	0.68	12.98	6.78	1.62
200	1	10.90	5.42	1.20	10.16	4.86	1.04	11.14	5.46	1.20
	2	12.14	6.34	1.54	9.40	4.24	0.88	12.38	6.50	1.56
	3	11.36	5.64	1.22	9.68	4.50	0.74	11.08	5.66	1.24
500	1	10.04	5.00	1.14	9.84	4.82	1.00	10.28	5.26	1.08
	2	10.80	5.54	1.44	9.52	4.74	1.02	10.86	5.82	1.42
	3	11.00	5.58	1.02	10.08	4.96	0.68	10.84	5.60	1.02
50	1	16.56	9.80	2.60	9.40	3.92	0.32	16.46	9.78	2.70
	2	21.92	13.68	4.66	8.60	2.90	0.14	21.66	13.36	4.76
	3	18.82	11.44	3.44	8.60	3.46	0.30	18.60	11.06	3.50
100	1	13.06	7.20	1.56	10.16	4.80	0.66	13.28	7.12	1.52
	2	15.96	9.24	2.20	9.04	3.40	0.48	15.96	9.24	2.36
	3	14.22	8.06	2.30	9.68	4.10	0.58	14.38	8.30	2.30
200	1	12.02	6.18	1.42	10.58	4.80	0.82	12.44	6.32	1.56
	2	13.50	6.92	1.80	9.00	4.08	0.52	13.34	6.96	1.72
	3	12.26	6.86	1.30	9.72	4.80	0.56	12.48	7.00	1.26
500	1	10.12	4.80	1.20	9.38	4.24	0.96	9.88	5.04	1.10
	2	12.08	6.34	1.20	9.76	4.62	0.64	12.10	6.34	1.38
	3	11.00	6.00	1.10	9.68	4.84	0.70	11.24	6.04	1.04
Power, $H_a^{\texttt{DPD}} : \lambda = (.05, .05, .05)$; `CH-0` (1st panel below); `CH-1` (2nd panel below)										
50	1	79.54	67.08	35.26	77.56	63.34	31.76	78.96	66.92	35.16
	2	64.46	48.44	24.30	82.54	71.86	50.42	64.16	49.84	22.08
	3	72.44	58.54	29.84	77.84	66.70	38.80	71.36	58.04	30.14
100	1	94.30	88.04	67.72	94.00	87.32	65.46	93.92	87.98	65.54
	2	85.78	75.86	43.16	93.68	89.14	72.50	85.62	76.22	44.40
	3	90.10	82.28	57.58	93.26	87.56	67.10	89.94	81.78	57.50
200	1	99.90	99.48	97.10	99.82	99.46	96.34	99.80	99.42	96.88
	2	99.12	97.82	88.42	99.68	99.32	96.66	99.04	97.66	88.12
	3	99.58	98.94	94.58	99.80	99.40	97.12	99.56	99.02	94.20
500	1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	2	100.00	100.00	99.98	100.00	100.00	100.00	100.00	100.00	100.00
	3	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
50	1	67.86	52.94	24.86	74.34	60.24	32.82	67.40	51.80	24.02
	2	56.30	39.72	14.12	80.36	71.98	50.06	55.50	38.96	12.90
	3	63.60	46.88	20.98	78.78	66.76	40.44	64.00	46.58	21.70
100	1	88.40	79.98	57.34	91.64	84.58	65.10	88.64	80.06	58.16
	2	78.62	66.98	38.48	93.78	89.48	74.62	79.08	67.06	36.82
	3	84.82	73.38	44.44	92.12	87.28	68.92	84.88	73.46	46.70
200	1	99.60	98.80	93.44	99.78	99.36	95.96	99.52	98.54	93.26
	2	97.34	93.92	79.36	99.44	98.86	96.38	97.26	93.56	79.40
	3	98.42	96.70	88.26	99.48	98.90	96.60	98.52	96.86	89.08
500	1	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	2	100.00	100.00	99.82	100.00	100.00	100.00	100.00	99.98	99.82
	3	100.00	100.00	99.98	100.00	100.00	100.00	100.00	100.00	100.00

Note: for dpg, 1 = normal, 2 = normal mixture, 3 = lognormal

Table 2b Empirical Size of Tests of $H_0^{\texttt{DPD}}: \lambda = 0$; `Group-I`, $\rho = 0.5$

		$T_{\texttt{DPD}}$			$T^{\diamond}_{\texttt{DPD}}$			$T^{\dagger}_{\texttt{DPD}}$		
n	dgp	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 3$, `CH-0` (1st panel below); `CH-2` (2nd panel below)										
50	1	14.18	7.26	1.52	11.30	4.76	0.48	14.08	7.46	1.58
	2	20.52	12.00	3.24	10.52	4.48	0.28	20.74	12.14	3.42
	3	15.80	9.22	2.36	10.50	4.26	0.46	16.14	9.14	2.26
100	1	12.10	6.44	1.36	10.54	5.38	0.86	12.28	6.58	1.26
	2	15.44	8.64	2.32	10.14	4.62	0.74	15.46	8.44	2.40
	3	14.52	7.96	2.08	11.24	5.46	0.66	14.60	8.26	2.16
200	1	12.04	6.06	1.54	10.90	5.46	1.38	11.96	6.28	1.60
	2	14.04	8.02	2.36	11.04	5.16	1.12	14.32	8.28	2.42
	3	11.88	6.38	1.40	10.18	5.12	0.94	11.80	6.44	1.38
500	1	11.20	5.82	1.12	10.88	5.70	0.88	11.32	5.84	1.10
	2	12.38	6.44	1.64	11.06	5.16	1.12	12.52	6.50	1.66
	3	11.44	5.82	1.24	10.86	5.14	0.88	11.54	6.00	1.18
	1	15.30	8.16	1.54	12.04	5.40	0.60	15.22	8.20	1.52
	2	18.96	11.00	3.14	9.82	3.62	0.30	18.96	10.98	2.98
	3	17.28	9.24	2.30	10.78	4.58	0.38	16.86	9.42	2.42
100	1	12.46	6.46	1.36	11.26	5.40	0.86	12.68	6.52	1.26
	2	16.00	8.84	2.50	10.24	4.82	0.78	16.18	8.98	2.48
	3	14.46	8.34	2.12	11.74	5.90	0.96	14.90	8.36	2.14
200	1	11.62	6.02	1.38	10.92	5.62	0.98	12.28	6.26	1.52
	2	13.02	6.74	1.78	9.78	4.74	0.84	13.68	7.22	1.72
	3	12.58	6.94	1.60	10.80	5.44	0.96	13.16	7.18	1.76
500	1	11.68	5.84	1.34	11.36	5.50	1.20	11.86	5.90	1.30
	2	11.68	5.78	1.38	9.88	4.78	1.02	11.80	5.90	1.46
	3	11.96	6.40	1.48	11.18	5.66	1.22	12.04	6.60	1.46
$T = 6$, `CH-0` (1st panel below); `CH-2` (2nd panel below)										
50	1	10.95	5.55	0.85	9.95	4.80	0.45	11.30	5.60	1.00
	2	14.65	7.95	1.40	9.15	3.20	0.35	14.35	7.50	1.40
	3	13.55	7.15	1.75	10.90	5.35	0.65	13.90	7.60	1.70
	1	11.60	5.70	1.25	11.05	5.30	1.10	11.80	5.70	1.20
	2	13.15	6.90	1.65	10.20	4.80	0.90	13.15	6.75	1.50
	3	12.45	5.60	1.50	10.30	4.35	0.90	11.90	5.40	1.50
200	1	10.65	6.00	1.45	10.60	5.30	1.40	10.95	5.70	1.55
	2	12.35	6.45	1.35	10.70	5.25	0.45	12.65	6.45	1.35
	3	10.10	5.40	1.75	9.00	4.85	1.40	9.75	5.60	1.75
500	1	11.05	5.85	1.40	10.80	5.75	1.45	11.25	5.80	1.55
	2	11.60	6.30	1.25	10.95	5.35	1.20	11.55	6.10	1.40
	3	10.55	5.15	0.90	10.20	4.70	0.80	10.30	4.90	0.95
	1	10.85	5.60	0.75	10.10	4.50	0.40	11.05	5.60	0.80
	2	13.80	7.75	1.20	8.20	3.60	0.00	14.10	7.85	1.30
	3	15.15	8.15	1.45	11.50	5.45	0.70	14.80	7.85	1.35
100	1	10.95	5.70	0.65	10.60	5.25	0.55	10.75	5.65	0.70
	2	13.85	8.10	2.00	11.55	5.70	0.75	14.00	7.90	2.15
	3	12.05	5.85	1.25	10.50	4.55	0.75	12.25	5.70	1.45
200	1	9.15	5.40	1.10	9.00	5.15	0.85	9.35	5.15	1.30
	2	11.50	5.15	1.00	9.75	4.25	0.65	11.35	5.70	1.10
	3	10.25	5.55	1.25	9.60	5.20	1.00	10.50	5.70	1.20
500	1	9.85	5.00	0.95	9.95	4.85	0.85	9.95	5.25	1.00
	2	11.65	5.55	0.85	10.90	4.65	0.35	11.45	5.45	0.65
	3	11.15	5.50	1.15	10.75	5.10	1.15	10.80	5.55	1.20

Note: for dpg, 1 = normal, 2 = normal mixture, 3 = lognormal

Table 2c Size-Adjusted Power of Tests of $H_0^{\mathtt{DPD}} : \lambda = 0$; `Group-I`, $\rho = 0.5$

		$T_{\mathtt{DPD}}$			$T^{\diamond}_{\mathtt{DPD}}$			$T^{\dagger}_{\mathtt{DPD}}$		
n	dgp	10%	5%	1%	10%	5%	1%	10%	5%	1%
$T = 3$, $H_a^{\mathtt{DPD}} : \lambda = (.05, .05, .05)'$, `CH-0` (1st panel below); `CH-2` (2nd panel below)										
50	1	59.82	46.54	25.02	57.32	45.42	22.90	58.92	46.72	24.46
	2	48.36	34.16	14.86	66.44	54.58	32.40	47.40	33.96	14.40
	3	54.62	40.20	16.68	61.98	50.00	24.54	55.18	40.46	16.36
100	1	69.88	56.76	31.34	69.48	55.16	30.88	69.72	56.52	32.46
	2	60.64	47.26	19.60	72.06	61.60	37.08	60.54	47.82	19.98
	3	64.48	50.74	26.08	69.44	56.46	35.08	63.36	50.78	25.44
200	1	61.34	46.90	21.04	61.20	46.80	21.38	60.62	46.84	21.48
	2	55.38	41.00	18.00	62.50	50.56	26.18	54.56	40.16	17.44
	3	60.64	48.02	23.50	63.68	50.68	27.06	60.78	48.12	23.38
500	1	97.56	95.06	86.28	97.58	95.14	86.10	97.58	94.84	86.46
	2	96.12	92.90	78.62	97.20	94.74	84.90	96.00	92.78	78.62
	3	96.88	94.18	83.80	97.16	94.96	86.20	96.74	93.76	83.22
	1	59.18	47.28	24.16	57.58	44.86	20.52	59.16	46.48	23.78
	2	51.02	35.48	13.20	67.60	56.32	36.08	50.60	35.48	13.24
	3	55.06	41.18	17.44	63.64	51.16	26.60	55.38	40.52	17.80
100	1	72.48	59.06	33.26	73.64	60.84	35.66	72.22	59.16	33.06
	2	63.32	47.32	22.80	76.66	66.86	43.26	62.62	46.90	22.12
	3	66.66	52.52	25.72	73.62	61.54	36.66	66.56	52.34	25.54
200	1	62.48	48.70	26.00	62.70	48.88	27.46	61.98	47.82	25.92
	2	60.30	46.80	23.62	67.56	55.86	33.86	60.42	46.54	23.64
	3	62.38	49.02	26.40	64.60	52.74	30.48	61.72	48.68	24.90
500	1	96.82	93.88	82.16	96.86	93.92	81.86	96.82	93.82	82.74
	2	95.60	91.78	78.30	96.68	94.08	83.72	95.46	92.22	77.70
	3	96.12	92.74	80.22	96.60	93.52	82.28	96.18	92.54	80.12
$T = 6$, $H_a^{\mathtt{DPD}} : \lambda = (.03, .03, .03)'$, `CH-0` (1st panel below); `CH-2` (2nd panel below)										
50	1	90.35	82.10	59.25	89.40	81.40	56.60	89.30	81.70	56.95
	2	79.70	68.75	42.25	91.45	85.70	65.15	79.00	69.30	39.55
	3	83.85	71.50	43.75	88.00	77.85	57.40	83.05	71.60	45.15
100	1	94.10	88.80	69.00	93.90	88.70	69.20	93.70	88.45	70.65
	2	90.60	83.75	57.00	95.00	90.75	76.20	91.10	82.80	58.40
	3	95.00	89.85	63.20	96.00	93.75	73.75	94.85	90.55	65.15
200	1	99.50	98.60	94.20	99.55	98.55	94.35	99.60	98.30	93.60
	2	99.00	97.70	90.60	99.35	98.90	95.70	99.10	97.40	90.35
	3	99.45	97.95	89.20	99.50	98.70	93.10	99.40	98.25	88.80
500	1	100.00	99.95	99.00	100.00	99.95	99.00	100.00	99.90	98.95
	2	100.00	99.85	98.00	100.00	99.95	98.75	100.00	99.70	97.55
	3	99.95	99.95	99.50	99.95	99.95	99.65	100.00	99.95	99.10
50	1	88.50	80.50	59.90	87.55	78.90	54.00	87.90	80.45	58.90
	2	78.10	65.40	42.05	88.65	82.55	70.50	76.75	63.70	38.30
	3	79.40	69.15	46.05	84.80	75.65	54.35	80.85	70.10	46.95
	1	97.05	93.20	81.80	97.15	93.90	83.30	96.90	92.80	81.95
	2	91.30	81.65	56.35	96.00	92.85	83.20	90.65	81.80	57.70
100	3	95.80	91.25	69.90	96.85	94.50	79.95	95.85	92.00	68.50
	1	99.60	98.85	94.60	99.60	98.90	95.20	99.60	99.05	93.95
200	2	98.85	97.90	91.90	99.30	98.70	96.40	98.95	97.80	90.10
	3	99.40	98.60	93.40	99.65	98.90	95.10	99.45	98.40	93.15
500	1	100.00	99.95	99.25	100.00	99.95	99.25	100.00	99.90	99.10
	2	100.00	99.90	99.05	100.00	99.95	99.75	99.95	99.85	99.10
	3	99.95	99.90	99.05	99.95	99.95	99.20	99.95	99.90	98.80

Note: for dpg, 1 = normal, 2 = normal mixture, 3 = lognormal

Table 3a Empirical Size of Tests of $H_0^{\texttt{SDPD4}}: \lambda_1 = \lambda_2 = 0$; `Queen`, $T = 3$

			$T_{\texttt{SDPD4}}$			$T^{\diamond}_{\texttt{SDPD4}}$			$T^{\dagger}_{\texttt{SDPD4}}$		
(ρ, λ_3)	n	dgp	10%	5%	1%	10%	5%	1%	10%	5%	1%
	Homoskedasticity, `CH-0`										
(-.5, .9)	50	1	12.24	6.00	1.18	10.14	4.30	0.58	11.34	6.00	1.32
		2	14.34	7.92	2.20	8.40	3.26	0.48	16.78	9.66	2.66
		3	12.54	6.56	1.30	8.90	3.88	0.46	13.98	7.46	1.94
	100	1	10.42	5.26	1.06	9.48	4.60	0.72	10.34	5.30	1.04
		2	13.36	7.22	2.00	9.50	4.74	0.68	13.94	7.78	1.90
		3	11.88	5.94	1.06	9.76	4.44	0.56	11.68	6.16	1.48
	200	1	10.16	4.74	1.08	9.50	4.28	0.92	11.06	6.06	1.34
		2	11.68	5.60	1.32	9.30	4.38	0.76	12.94	6.76	1.52
		3	11.16	6.08	1.46	10.12	5.02	0.88	12.22	6.56	1.60
(.5, -.9)	50	1	11.26	5.58	1.24	9.68	4.12	0.74	10.80	5.86	1.04
		2	15.30	8.38	2.40	9.24	3.90	0.34	15.40	8.20	2.20
		3	12.18	6.64	1.42	9.78	4.52	0.56	12.32	6.64	1.32
	100	1	11.34	5.48	1.06	10.04	4.82	0.80	10.80	5.72	1.02
		2	12.88	7.14	1.60	9.42	4.26	0.40	13.22	7.38	1.66
		3	10.82	5.60	1.10	8.90	3.98	0.72	12.06	6.10	1.42
	200	1	9.62	4.90	0.92	9.18	4.48	0.68	10.16	4.88	0.84
		2	11.54	6.22	1.22	10.06	4.62	0.64	11.90	6.56	1.50
		3	10.84	5.68	1.18	9.72	4.72	0.84	11.64	5.96	1.08
(-.9, .9)	50	1	12.32	5.38	0.80	9.50	3.66	0.42	10.18	4.38	0.62
		2	16.32	8.58	2.02	8.38	3.10	0.16	12.66	6.70	1.12
		3	13.74	6.90	1.34	8.96	3.56	0.20	11.96	5.80	1.14
	100	1	10.42	5.02	0.70	9.28	4.12	0.42	9.02	4.12	0.50
		2	16.40	9.14	2.02	10.96	4.56	0.44	10.90	5.32	1.06
		3	13.22	6.98	0.96	10.16	4.44	0.38	10.54	4.84	0.76
	200	1	9.94	4.64	0.80	9.36	4.22	0.64	9.88	4.70	0.42
		2	13.16	6.76	1.46	10.28	4.54	0.62	10.38	5.40	0.70
		3	11.80	6.08	0.98	10.44	4.78	0.58	11.06	5.18	1.00
	Heteroskedasticity, `CH-1`										
(-.5, .9)	50	1	14.02	7.58	1.98	9.44	4.12	0.50	14.24	8.04	1.94
		2	17.56	10.78	3.28	9.34	3.40	0.28	18.14	11.06	3.80
		3	15.46	9.20	2.74	8.70	3.96	0.64	16.88	10.36	2.96
	100	1	11.88	5.90	1.26	9.94	4.64	0.64	12.06	6.52	1.36
		2	14.60	8.32	2.48	9.88	4.70	0.66	15.82	8.94	2.84
		3	13.32	7.12	2.00	10.38	4.38	0.92	13.74	7.64	1.78
	200	1	10.78	5.40	1.08	9.50	4.44	0.64	12.12	6.40	1.32
		2	12.66	7.24	1.70	9.24	4.36	0.68	13.74	7.78	2.02
		3	12.08	6.02	0.96	9.80	4.30	0.44	13.20	7.30	1.68
(.5, -.9)	50	1	14.12	8.18	2.12	9.74	4.28	0.42	14.74	7.88	2.16
		2	19.92	12.14	4.40	9.46	3.80	0.50	20.34	12.32	4.28
		3	16.24	9.62	2.72	9.50	4.04	0.48	15.16	8.74	2.14
	100	1	11.78	6.54	1.28	10.20	4.78	0.64	12.34	6.38	1.34
		2	14.90	8.42	2.04	9.76	3.96	0.54	15.08	8.00	2.12
		3	12.82	6.86	1.68	9.70	4.46	0.64	13.48	6.86	1.74
	200	1	10.38	5.62	1.32	9.52	4.78	0.94	10.86	5.80	1.36
		2	11.94	6.26	1.46	8.84	3.82	0.58	13.00	6.56	1.58
		3	11.86	6.28	1.46	9.80	4.66	0.76	12.42	6.56	1.56

Note: for dpg, 1 = normal, 2 = normal mixture, 3 = lognormal

Table 3b Empirical Size of Tests of $H_0^{\texttt{SDPD4}}: \lambda_1 = \lambda_2 = 0$; $(\rho, \lambda_3) = (.5, .3)$, $T = 3$

			$T_{\texttt{SDPD4}}$			$T^{\diamond}_{\texttt{SDPD4}}$			$T^{\dagger}_{\texttt{SDPD4}}$		
CH	n	dgp	10%	5%	1%	10%	5%	1%	10%	5%	1%
			Group-I								
CH-3	50	1	11.54	6.06	1.88	6.52	2.98	0.30	14.46	8.30	2.18
		2	15.76	8.78	2.60	6.30	2.28	0.10	17.90	10.60	3.18
		3	13.38	7.48	2.16	6.32	2.26	0.22	15.82	8.94	2.34
	100	1	15.12	9.08	2.48	11.32	5.80	0.76	12.76	7.52	1.88
		2	19.60	12.30	4.52	10.84	4.50	0.46	16.16	9.66	3.22
		3	17.24	11.00	3.56	11.38	5.82	0.78	13.82	7.66	2.00
	200	1	10.50	5.30	1.12	8.38	3.86	0.56	10.68	5.40	1.16
		2	12.84	7.08	1.70	7.90	3.20	0.48	14.30	7.92	2.00
		3	11.88	6.24	1.50	8.56	3.30	0.40	12.80	6.56	1.80
	500	1	11.12	6.32	1.40	10.36	5.48	0.92	11.22	5.76	1.20
		2	11.44	6.34	1.52	8.90	3.98	0.58	12.10	6.44	1.50
		3	11.68	6.24	1.78	9.96	4.70	0.92	10.40	5.28	1.28
			Group-II								
CH-1	50	1	13.02	6.74	1.20	9.82	4.30	0.62	12.78	6.26	0.88
		2	16.90	9.60	2.70	9.06	3.40	0.34	15.64	8.58	2.42
		3	15.18	7.82	1.58	9.00	3.76	0.32	13.46	7.02	1.36
	100	1	10.20	4.80	0.82	8.58	3.48	0.42	10.62	5.20	1.10
		2	13.88	7.76	2.02	8.90	3.80	0.44	13.08	6.92	1.74
		3	11.00	5.48	1.08	8.26	3.56	0.42	11.42	5.32	1.14
	200	1	9.98	5.32	1.04	9.16	4.60	0.82	10.46	5.44	1.20
		2	11.28	5.86	1.30	9.02	3.96	0.60	11.22	5.86	1.56
		3	10.46	5.14	1.14	9.04	3.88	0.62	11.16	5.46	1.14
	500	1	8.80	4.16	0.58	8.44	3.88	0.46	10.44	4.62	0.84
		2	9.76	4.82	0.96	8.62	4.10	0.58	10.76	5.38	0.96
		3	9.36	4.32	0.84	8.28	3.54	0.50	10.44	5.18	1.12
CH-2	50	1	12.92	6.44	1.28	11.70	5.06	0.82	10.40	5.18	0.96
		2	16.12	9.06	2.28	9.84	4.14	0.44	13.76	7.00	1.64
		3	14.52	8.32	1.82	11.44	5.50	0.56	12.30	6.32	1.32
	100	1	12.66	6.76	1.44	11.86	5.90	1.18	11.06	5.82	1.18
		2	15.12	8.70	2.36	11.60	5.66	0.86	12.32	6.48	1.54
		3	12.40	6.26	1.32	10.60	4.94	0.86	10.42	5.32	1.06
	200	1	11.86	6.30	1.24	11.40	5.98	1.16	10.10	4.88	0.68
		2	13.72	7.40	1.80	11.28	5.54	0.98	10.92	5.54	1.06
		3	12.48	6.86	1.42	11.60	6.12	1.20	10.72	5.30	1.16
	500	1	13.02	7.24	1.80	12.92	7.10	1.76	10.12	4.72	1.00
		2	13.22	7.46	2.20	12.48	6.66	1.78	10.52	5.44	0.86
		3	13.54	7.16	1.48	13.04	6.76	1.34	9.66	5.22	0.90
CH-3	50	1	21.82	14.70	5.70	13.42	6.22	0.68	18.92	11.40	3.24
		2	26.92	18.62	8.60	9.46	3.90	0.26	22.56	14.38	5.36
		3	22.66	15.02	5.94	11.10	4.74	0.44	19.64	11.86	4.22
	100	1	16.88	11.10	4.02	11.64	6.48	0.98	16.30	9.50	2.74
		2	21.88	14.62	5.92	9.76	4.28	0.42	19.24	12.18	4.26
		3	20.18	13.46	5.26	12.28	5.30	0.80	16.90	10.14	3.36
	200	1	18.68	10.84	3.06	17.96	10.08	2.24	11.68	6.04	1.24
		2	20.62	13.12	4.42	15.92	8.64	1.60	13.12	7.38	1.90
		3	19.30	11.62	3.40	16.42	9.10	1.76	11.44	6.20	1.60
	500	1	27.78	18.32	6.76	27.52	17.86	6.54	10.72	5.42	1.52
		2	28.76	17.96	5.84	25.94	15.12	3.86	11.50	5.96	1.40
		3	28.10	19.00	7.02	26.74	17.58	5.52	11.46	5.68	1.14

Note: for dpg, 1 = normal, 2 = normal mixture, 3 = lognormal