Specification Search for Growth Model: A Dynamic Space -Time Framework

Monalisa Sen *
University of Illinois at Urbana Champaign
Job Market Paper

November 9, 2011

ABSTRACT

In this paper I propose adjusted Rao's score (RS) tests that are robust under misspecification for a very general dynamic panel model with cross-sectional dependence and use them for specification search of growth models. In growth theory different kinds of econometric models have been proposed based on economic theory and the subjective beliefs of researchers, - including simple cross-sectional regression, panel data, time series and more recently many types of spatial models. Unfortunately the estimate of growth convergence rate under these different model frameworks vary wildly, even when the same dataset is used. Thus, the question is which model is most appropriate? I use my proposed tests to address this problem and conduct the specification search in multiple directions to understand the underlying data generating process (DGP). Unlike the available tests, these proposed test statistics unravel the interrelation/dependencies among the model parameters and thus make themselves amenable for analysis of misspecification, which is concept-wise similar to analysis of variance. I use the data of 91 non-oil countries over a period of 35 years (1961- 1995) from the Penn World Table, for the specification search. Using the proposed test statistics, I find that heterogeneity, time dynamics and indirect cross-sectional dependence contribute most to the total misspecification than other forms of departures from a simple panel model for this dataset. A very elegant feature of my proposed tests is that they do not require estimation of nuisance parameters, unlike existing tests. Thus the proposed test statistics can identify the underlying DGP without any apriori complex estimation. The extensive simulation study show good finite sample performance of my proposed tests in contrast to other existing procedures. The formulation of these test statistics are quite general and are applicable to many other econometric models for specification search.

Keywords: Growth Convergence, Rao's score tests, Panel Data Methods, Spatial Dependence.
JEL classification: C01, C12, C18, C33, O47
I am grateful to my advisor, Anil K. Bera, for his continuous encouragement, advice and generous support in this project. I am also thankful to my thesis committee members Geoffrey Hewings, Daniel McMillen, and Xiaofeng Shao for their fruitful advice at the crucial stages of my work. My thanks also goes to Dan Bernhardt, Ingmar Prucha and UIUC seminar participants for helpful discussions. I also thank Paul Boltz Fellowship, awarded by the University of Illinois for supporting my research during 2011. Any remaining errors and omission are my responsibility.
* *Address*: Dept. of Economics, University of Illinois at Urbana-Champaign, 214 David Kinley Hall, 1407 W. Gregory Drive, Urbana, IL 61801. Email: mde2@illinois.edu. Comments are most welcome.

1 Introduction

"It isn't what we don't know that kills us. It's what we know that ain't so." –Mark Twain

Understanding the nation's growth is one of the oldest and most important research agendas in economics. At the same time, the empirical study of economic growth occupies a position that is notably uneasy. Rodriguez and Rodrik (2001) begin their skeptical critique of evidence on trade policy and growth with the above quote which they use to point out the difficulty in identifying the salient determinants of growth. Quite generally, one such difficult issues is the basic econometric specification of the growth models, i.e., absence of consensus regarding the salient features of the underlying data generating process (DGP). There is a vast literature on econometric issues that arise due to different presumptions on the structure of the DGP that appear in growth analyses. The choices of method involve significant trade-offs, which depend both on statistical considerations and on economic context. Inspite of the vast literature on this particular issue of econometric specification of the models, it has always been difficult to identify the structure of DGP. In this paper, I address this point. I formulate new diagnostic tests that take into account of misspecification in multiple directions. In particular, I propose adjusted Rao's score (RS) test statistics under a dynamic panel spatial model framework, which are robust under misspecification. I use the proposed tests for *specification search in multiple directions*, without any complex estimation of the nuisance parameters. The proposed test statistics can assist a researcher to revise his/her model towards appropriate direction(s) for better understanding of the growth behavior and thereby suggest suitable policy reforms.

Research on convergence proceeded through several stages and also witnessed the use of different methodologies. However, the correspondence between the convergence definitions (like β-convergence, σ-convergence, conditional convergence, and so on) and the methodologies used are not unique. For example, cross-section, panel and time series (in part) approaches have been used to study β. These approaches have generally dealt with convergence across economies and in terms of per capita income level. The cross-sectional approach is popular to study σ-convergence, while time series methodologies are implemented to investigate convergence both within and across economies. More recently, various spatial approaches has been adopted to model the technological spillover and interdependencies of economies, both in cross-sectional and panel framework. Each of these methodologies has its own benefits and drawbacks, even though it may be used to study a particular aspect of convergence. For example, the use of panel data to study β-convergence, is likely to increase efficiency and allow richer models in the presence of parameter heterogeneity. Thus there is a trade-off between robustness and efficiency with each of the chosen methods. The scientific solution would be to base the choice of estimation method on a context-specific loss function. This is clearly a very difficult task. Thus the crucial question is which model/models do the data confirm with?

This paper provides a solution to this difficult problem in the context of growth model, i.e., unravel the DGP without any subjective preference, so that the researcher can choose a suitable model to understand the underlying growth behavior. As mentioned earlier, there has been many studies that have considered only cross-section, or time series, or panel, or spatial model methods. In this paper I consider the dynamic panel spatial model framework which is a generalization of all these piece-wise models and propose test statistics to understand which kind of departures are actually present in the underlying dataset. I start with a small model (simple panel model under joint null hypothesis) and then check whether specific departures (like time dynamics, serial correlation of errors, individual effects, different forms of spatial dependencies) from this starting model are supported or rejected by the data. Using Bera and Yoon (1993) test principle, I propose new adjusted Rao's score (RS) test statistics for each of the parameters, after taking into account the possible presence of all other forms of departures. Unlike the existing conditional tests, the proposed methodology takes care of the possible presence of all the nuisance parameters through their respective Fisher-Rao score evaluated under *joint null*, and thus requiring estimation of the simplest model. Using these proposed test statistics I also show how some existing models are potentially misspecified.

The main contributions of this paper are thus twofold: (i) development of *six* new RS test statistics robust under local misspecification, i.e., adjusted RS for time dynamics, random effects, serial correlation, space-time dynamics, spatial lag and spatial dependence parameters, where each of them is robust to the presence of all the other departures (nuisance parameters). The proposed test statistics do not require any estimation of the nuisance parameters and thus, are computationally simple and easily amenable for misspecification analysis. (ii) Using these proposed test statistics, I address the empirical question of specification search, i.e., which model/models do the underlying data for growth models confirm with? Thus, using my proposed tests I come up with a proper model for the growth analysis.

The plan of the rest of the paper is as follows. The next section provides a brief review of the existing models used for growth convergence. I provide the details of our model framework and the regularity conditions in Section 3. In Section 4, I derive the new diagnostic tests which take account of misspecification in multiple directions. After reviewing the data set in Section 5, I discuss how the proposed test statistics can unravel the dependencies of the underlying DGP without any complex estimation of the model itself. Thus I propose an appropriate growth model that capture the salient features of the data. In Section 6, I conduct finite sample study to evaluate the performance of the suggested and some available tests, and Section 7 concludes the paper.

2 A Brief Literature Review

The literature on growth convergence initiated by the seminal papers of Solow (1956) and Swan (1956) is vast and it reached the 'formal specification' stage with the influential work of Barro and Salai-Martin (1992) (henceforth BS) and Mankiw, Romer and Weil (1992), (henceforth MRW), which derived the regression specification from the neoclassical growth model. MRW is based on original Solow-Swan model, and BS on Cass-Koopmans' (1965) optimal savings model. Both papers derive the law of motion of capital and income around the steady state and then translate that into an estimable *cross-sectional regression* equation. Similar results on conditional convergence across countries are presented in Holtz-Eakin (1993), Sala-i-Martin (1996) and many others. One of the crucial assumption of these cross-sectional models is that the differences in initial unobserved technology diffusion is considered to be a part of error terms. This assumption makes their equations estimable by ordinary least square (OLS) method. Thus the cross-section approach to convergence encountered some important limitations. Temple (1998) discussed the influence of possible measurement errors on the results of MRW. The basic limitation lies in the fact that having just one data for a country provides a weak basis for estimation of the convergence, which refers primarily to a within-country process. There is too much heterogeneity across countries to validate the assumption that cross country data can be treated as multiple data of the same country. Thus, the convergence research gradually moved to other approaches like time-series and panel methods.

Lee, Pesaran and Smith (1992), Quah (1992), Binder and Pesaran (1999) support for *time series regression* for each country separately to analyze the conditional convergence hypothesis. In simple terms, convergence using time-series approach, implies whether income of a specific country has unit root or not. They argue that standard cross-section methods throw away useful information which can be taken care by analyzing each country separately. Moreover, time series analysis has been applied to investigate across convergence too, see for instance Quah (1992), Bernard and Durlauf (1995) and Evans and Karras (1996). Broadly speaking the time series analysis supports a variant of conditional convergence hypothesis and thus results are not much different from those implied by cross-section methods.

One of the crucial limitation of the cross-section approach is that it cannot capture the technological diffusion and capital deepening process, which are vital for income convergence across countries. Thus many researchers used *panel methods* to capture such technology diffusion by introducing individual effects in the model. However, there are many ways to model the *country-specific effect*. For instance, Islam (1995) strongly supports fixed effect estimation due to the assumption of correlation of unobserved technology diffusion with the regressors. The key strength of this method is that it takes care of one form of heterogeneity: any omitted variables that are constant over time will not bias the estimates, even when the omitted variables are correlated with the explanatory variables.

There are, however, some concerns about fixed effect specification. For instance, sometimes a variable of interest is measured at only one point in time, and even if the variables are measured at more frequent intervals, they are sometimes highly persistent. In that case the within-country variation is unlikely to be informative. Too often researchers use fixed effects to analyze the effects of variables that are fairly constant over time, or that affect growth only with a long time lag. Standard transformation like first differences or "within groups" transformations are likely to exacerbate the problem of measurement errors. They lead to large reduction in precision of the parameter estimates since the between-country variation is thrown away. Koop and Steel (2000) argues that much of variation in efficiency level occurs *between* rather than *within* countries. Thus, a random effects generalized least square (GLS) estimator will be more efficient than within-country estimator when the random effects assumptions are appropriate. Durlauf and Quah (1999) point out that the individual effects are of fundamental interest to growth economists as they appear to be the key source of persistent income differences. Thus they advocate for modeling the heterogeneity in the model rather than finding the ways to eliminate its effects. In this paper, I adopt a random effects model as I intend to *test* the significance of individual effects in the presence of time dynamics and spatial dependencies, rather than just treating them as the nuisance parameters, as is done in fixed effects model.

Recently, many researchers are using *spatial models* to analyze growth convergence. It is a known fact that the economies are assumed to be independent in the neoclassical growth theory. However, with globalization, technological advances in one economy are transmitted to other economies. Thus, the closed independent economy assumption are not valid, and one needs to take into account the possible spatial correlation, both in cross-sectional and panel data settings. From statistical point of view, ignoring the presence of spatial dependence leads to unreliable inference. In recent years many researchers have used spatial methods to capture such technology interdependence and knowledge spillover effects. The main idea is to capture the impact of cross-country spillovers on growth process. There are many ways to measure this interdependence. One of the most common way is to express the aggregate level of technology of any country to be dependent on the stock of knowledge/capital of its neighbors or trading partners by using geographic and economic distances. For instance, Conley and Ligon (2002), Ertur, Gallo and Baumont (2006), Ertur and Koch (2007), Yu and Lee (2009) and many others, have used spatial approach to analyze growth convergence.

Each method, as I discussed, has its own merits and drawbacks. It is evident from the discussion that the convergence research has not produced any concrete consensus. Given the differences in approach, sample, data, model, estimation technique, etc., absence of consensus is not surprising though. The crucial issue is to find a good approximation to DGP, and to achieve this objective, I start out with a general model in the following section.

3 The Model Setup

The model setup is the combination of all the piecewise frameworks I discussed earlier:

$$y_{it} = \gamma y_{it-1} + \tau \sum_{j=1}^{N} m_{ij} y_{jt} + \delta \sum_{j=1}^{N} m_{ij} y_{jt-1} + X_{it}\beta + u_{it} \tag{1}$$

$$u_{it} = \mu_i + \epsilon_{it} \tag{2}$$

$$\epsilon_{it} = \lambda \sum_{j=1}^{N} w_{ij}\epsilon_{jt} + v_{it} \tag{3}$$

$$v_{it} = \rho v_{it-1} + e_{it}, \text{where } e_{it} \sim IIDN(0, \sigma_e^2) \tag{4}$$

for $i = 1, 2, \ldots, N; t = 1, 2, \ldots, T$. Here y_{it} is the observation for i^{th} location/unit at t^{th} time, X_{it} denotes the observations on non-stochastic regressors and e_{it} is the regression disturbance. Spatial dependence is captured by the weight matrices $M = (m_{ij})$ and $W = (w_{ij})$. Here m_{ij} and w_{ij} are the (i, j) th element of weight matrices M and W respectively, which capture the interdependence of income and unobserved error terms between the country i and j.The matrices M and W are each row-standardized and the diagonal elements are set to zero. In this model framework, time dynamics (γ), random effects (μ_i) with $\mu_i \sim IID(0, \sigma_\mu)$, serial correlation ($\rho$), space recursive ($\delta$), spatial lag dependence (τ) and spatial error dependence (λ) are considered.

In matrix form, the equations (1) - (4) can be written compactly as

$$y = \tau(I_T \otimes M)y + [(\gamma + \delta M) \otimes I_T]ly + X\beta + u, \tag{5}$$

where y is of dimension $NT \times 1, X$ is $NT \times K, \beta$ is $k \times 1$, u is $NT \times 1$, I_T is an identity matrix of dimension $T \times T$ and $\otimes$ denotes Kronecker product. Here l is the lag operator, X is assumed to be of full column rank and its elements are bounded in absolute value. The disturbance term can be expressed as

$$u = (\iota_T \otimes I_N)\mu + (I_T \otimes B^{-1})v. \tag{6}$$

Here $B = (I_N - \lambda W)$ and ι_T is vector of ones of dimension T. Under this setup, the variance-covariance matrix of u is given by

$$\Omega = \sigma_{\mu^2}[J_T \otimes I_N] + [V \otimes (B'B)^{-1}], \tag{7}$$

where J_T is a matrix of ones of dimension $T \times T$, and V is the familiar $T \times T$ variance -covariance matrix for AR (1) process in equation (12, i.e.,

$$V = E(v'v) = [\frac{1}{1-\rho^2}V_1] \otimes \sigma_e^2 I_N = V_\rho \otimes \sigma_e^2 I_N, \tag{8}$$

with

$$V_1 = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \dots & \dots & 1 \end{bmatrix},$$

and $V_\rho = \frac{1}{1-\rho^2} V_1$.

The log-likelihood function of the above model can be written as:

$$L = \frac{-NT}{2} ln 2\pi - \frac{1}{2}\ln|\Omega| + T\ln|A| - \frac{1}{2}[(I_T \otimes A)y - [(\gamma+\delta M)\otimes I_T]ly - X\beta]'\Omega^{-1}[(I_T\otimes A)y - [(\gamma+\delta M)\otimes I_T]ly - X\beta] \tag{9}$$

where $A = (I_N - \tau M)$. Following Sen and Bera (2011), I can write

$$\frac{1}{2}\ln|\Omega| = -\frac{N}{2}\ln(1-\rho^2) + \frac{1}{2}\ln|d^2(1-\rho)^2\phi I_N + (B'B)^{-1}| + \frac{NT}{2}\ln\sigma_e^2 - (T-1)\ln|B|,$$

where $d^2 = \alpha^2 + (T-1)$, $\alpha = \sqrt{\frac{1+\rho}{1-\rho}}$ and $\phi = \frac{\sigma_\mu^2}{\sigma_e^2}$. Substituting $\frac{1}{2}\ln|\Omega|$ in L , I obtain

$$L = \frac{-NT}{2} ln 2\pi + \frac{N}{2}\ln(1-\rho^2) - \frac{1}{2}\ln|d^2(1-\rho)^2\phi I_N + (B'B)^{-1}| - \frac{NT}{2}\ln\sigma_e^2 + (T-1)\ln|B| + T\ln|A| \\ - \frac{1}{2}[(I_T\otimes A)y - [(\gamma+\delta M)\otimes I_T]ly - X\beta]'\Omega^{-1}[(I_T\otimes A)y - [(\gamma+\delta M)\otimes I_T]ly - X\beta] \tag{10}$$

The likelihood function in equation (10), will be used to derive the test statistics in the next section. Now I state the assumptions, needed for the validity of the asymptotic properties.

Assumption 1.

(i) W and M are row-standardized weight matrices whose diagonal elements are zero.

(ii) W and M are uniformly bounded in row and column sums in absolute value and $(I - \lambda W)^{-1}$ and $(I - \tau M)^{-1}$ are also uniformly bounded.

(iii) $A_n = (I - \tau M)^{-1}(\gamma I + \delta M)$ is also uniformly bounded in row and column sums in absolute value.

Assumption 2. The disturbances $e_{it}, i = 1, 2, \dots, N$ and $t = 1, 2, \dots, T$, are *i.i.d.* across i and t with zero mean, variance σ_e^2 and $E|e_{it}|^{4+\eta} < \infty$, for some $\eta > 0$.

Assumption 3. The element of X_{Nt} are nonstochastic and bounded uniformly (BU) in n and T. Also, $\lim_{T\to\infty} \frac{1}{NT}\sum_{t=1}^{T} X'_{Nt}X_{Nt}$ exists and is nonsingular.

Assumption 4. N is a nondecreasing function of T and T goes to infinity.

Assumption 1 is standard assumption in spatial analysis and boundedness condition on W, M and

$(I - \lambda W)$, $(I - \tau M)$ and A_n, limit the spatial correlation of manageable degree. *Assumption 2* provides regularity assumptions on for e_{it}. When exogenous variables X_{Nt} are included in the model, it is convenient to assume that they are uniformly bounded as in *Assumption 3*. Lastly, if *Assumption 4* holds then we can say that $N, T \to \infty$ simultaneously.

4 Derivation of the Test Statistics

The full model (1) - (4) has the following: linear regression coefficients and innovation variance (β, σ_e^2), time dynamics γ, random effects σ_μ^2, time-series correlation ρ, space-time dynamics parameter δ, spatial error dependence λ, and spatial lag dependence τ. The full parameter vector will be denoted by $\theta = (\beta', \sigma_e^2, \gamma, \sigma_\mu^2, \rho, \delta, \lambda, \tau)'$. I am interested in testing significance of last six parameters individually in the possible presence of the rest. For example, in order to detect the time- dynamics I would test, say, $H_o^b : \gamma = 0$ in presence of $\phi = (\sigma_\mu^2, \rho, \delta, \tau, \lambda)'$. The usual practice is using likelihood ratio test and conditional RS tests. However, those tests require estimation of both γ and ϕ (or ϕ alone) along with (β, σ_e^2). Instead, in this paper I contribute to the existing literature by developing adjusted RS tests for specification search in dynamic panel spatial framework by using Bera and Yoon (1993) test principle, which requires estimation of the simplest model under joint null of no misspecification.

All the proposed adjusted tests are based on the joint null hypothesis (of no misspecification), i.e., $H_0^a : \gamma = \sigma_\mu^2 = \rho = \delta = \lambda = \tau = 0$. Thus under H_o^a, the parameter vector is $\theta_o = (\beta', \sigma_e^2, 0, 0, 0, 0, 0, 0)'$. The proposed tests can take care of the possible presence of all the nuisance parameters indirectly through their respective Fisher-Rao score evaluated under the joint null. Due to this estimation simplicity, the suggested tests are more amenable to use by empirical researchers than the LR or conditional RS tests.

4.1 Bera and Yoon Test Principle

Consider a general model represented by the log-likelihood $L(\omega, \psi, \phi)$ where the parameters ω, ψ and ϕ are, respectively, $(p \times 1), (r \times 1)$ and $(s \times 1)$ vectors. Here I assume that underlying density function satisfies the regularity conditions, as stated in Serfling (1980), Lehmann and Romano (2005), for the MLE to have asymptotic Gaussian distribution. Suppose a researcher sets $\phi = \phi_0$ and tests $H_0 : \psi = \psi_0$ using the log-likelihood function $L_1(\omega, \psi) = L(\omega, \psi, \phi_0)$, where ψ_0 and ϕ_0 are known. The RS statistic for testing H_0 in $L_1(\omega, \psi)$ will be denoted by RS_ψ. Let us denote $\theta = (\omega', \psi', \phi')'_{(p+r+s)\times 1}$ and $\tilde{\theta} = (\tilde{\omega}', \psi_0', \phi_0')'_{(p+r+s)\times 1}$, where $\tilde{\omega}$ is the ML estimator of ω under $\psi = \psi_0$ and $\phi = \phi_0$. I define the score vector and the information matrix, respectively, as

$$d_a(\theta) = \frac{\partial L(\theta)}{\partial a} \qquad \text{and} \qquad J(\theta) = -E[\frac{1}{n}\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}] = \begin{bmatrix} J_\omega & J_{\omega\psi} & J_{\omega\phi} \\ J_{\psi\omega} & J_\psi & J_{\psi\phi} \\ J_{\phi\omega} & J_{\phi\psi} & J_\phi \end{bmatrix} \tag{11}$$

where $a = (\omega, \psi, \phi)$ and n is the sample size. If $L_1(\omega, \psi)$ were the true model, then it is well known that under $H_0 : \psi = \psi_0$

$$RS_\psi = \frac{1}{n} d_\psi(\tilde{\theta}) J_{\psi.\omega}(\tilde{\theta})^{-1} d_\psi(\tilde{\theta})' \to \chi^2_r(0) \tag{12}$$

where $J_{\psi.\omega}(\tilde{\theta}) = J_\psi - J_{\psi\omega} J_\omega^{-1} J_{\omega\psi}$.

Under local alternative $H_1 : \psi = \psi_0 + \frac{\zeta}{\sqrt{n}}$, $RS_\psi \to \chi^2_r(\lambda_1)$, where the non-centrality parameter $\lambda_1 \equiv \lambda_1(\zeta) = \zeta' J_{\psi.\omega}\zeta$. Given this setup, i.e., under no misspecification, asymptotically the test will have the correct size and locally optimal. Now suppose that the true log-likelihood function is $L_2(\omega, \phi)$ so that the considered alternative $L_1(\omega, \psi)$ is (completely) misspecified. Using the local misspecification $\phi = \phi_0 + \frac{\delta}{\sqrt{n}}$, Davidson and MacKinnon (1987) and Saikkonen (1989) derived the asymptotic distribution of RS_ψ under $L_2(\omega, \phi)$ as $RS_\psi \to \chi^2_r(\lambda_2)$, where the non-centrality parameter $\lambda_2(\delta) = \delta' J_{\phi\psi.\omega} J_{\psi.\omega}^{-1} J_{\psi\phi.\omega}\delta$ with $J_{\psi\phi.\omega} = J_{\phi\psi} - J_{\phi\omega} J_\omega^{-1} J_{\omega\psi}$. Owing to the presence of this non-centrality parameter λ_2, RS_ψ will reject the true null hypothesis $H_0 : \psi = \psi_0$ more often, i.e., the test will have excessive size. Here the crucial term is $J_{\phi\psi.\omega}$ which can be interpreted as partial covariance between the score vectors d_ϕ and d_ψ after eliminating the linear effect of d_ω on d_ϕ and d_ψ. If $J_{\psi\phi.\omega} = 0$, then asymptotically the local presence of ϕ has no effect on RS_ψ. Bera and Yoon (1993) suggested a modification to RS_ψ to overcome this problem of over-rejection, so that the resulting test is valid under the local presence of ϕ. The modified statistic is given by

$$RS^*_\psi = \frac{1}{N}[d_\psi(\tilde{\theta}) - J_{\psi\phi.\omega}(\tilde{\theta}) J^{-1}_{\phi.\omega}(\tilde{\theta}) d_\phi(\tilde{\theta})]' [J_{\psi.\omega}(\tilde{\theta}) - J_{\psi\phi.\omega}(\tilde{\theta}) J^{-1}_{\phi.\omega}(\tilde{\theta}) J_{\phi\psi.\omega}(\tilde{\theta})]^{-1} [d_\psi(\tilde{\theta}) - J_{\psi\phi.\omega}(\tilde{\theta}) J^{-1}_{\phi.\omega}(\tilde{\theta}) d_\phi(\tilde{\theta})]'. \tag{13}$$

This new test essentially adjusts the mean and variance of the standard RS statistics RS_ψ, and, under $H_o : \psi = \psi_o$

$$RS^*_\psi \to \chi^2_r(0) \tag{14}$$

while under $H_1 : \psi = \psi_0 + \frac{\zeta}{n}$,

$$RS^*_\psi \to \chi^2_r(\lambda_3) \tag{15}$$

where $\lambda_3 = \zeta'(J_{\psi.\omega} - J_{\psi\phi.\omega} J^{-1}_{\psi.\omega} J_{\phi\psi.\omega})\zeta$. Note the results in (14) and (15) are valid both under presence or absence of local misspecification, since the asymptotic distribution of RS^*_ψ is unaffected by the local departure of ϕ from ϕ_0.

BY shows that for local misspecification the adjusted test is asymptotically equivalent to Neyman's $C(\alpha)$ test and thus shares its optimal properties. Three observations are worth noting regarding RS^*_ψ. First, RS^*_ψ requires estimation only under the joint null, namely $\psi = \psi_0$ and $\phi = \phi_0$. That means, in most cases, as we will see later, we can conduct our tests based on only OLS residuals. Given the full specification of the model $L(\omega, \psi, \phi)$, it is of course possible to derive RS test for $\psi = \psi_0$ after estimating ϕ (and ω) by MLE, which are generally referred to as conditional tests. However, ML estimation of ϕ could be difficult in some instances. Second, when $J_{\psi\phi.\omega} = 0, RS^*_\psi = RS_\psi$, which is a simple condition to check. If this condition is true, RS_ψ is an asymptotically valid test in the local presence of ϕ. Finally, let $RS_{\psi\phi}$ denote the joint RS test statistic for testing hypothesis of the form $H_0 : \psi = \psi_0$ and $\phi = \phi_0$ using the alternative model $L(\omega, \psi, \phi)$. Then it be shown that [for a proof see Bera, Bilias and Yoon (2007), Bera, Montes-Rojas and Sosa-Escudero (2009)]

$$RS_{\psi\phi} = RS^*_\psi + RS_\phi = RS^*_\phi + RS_\psi, \tag{16}$$

where RS_ϕ and RS^*_ϕ are, respectively, the counterparts of RS_ψ and RS^*_ψ for testing $H_0 : \phi = \phi_0$. This is a very important identity since it implies that a joint RS test for two parameter vectors ψ and ϕ can be decomposed into sum of two orthogonal components: (i) the adjusted statistic for one parameter vector and (ii) (unadjusted) marginal test statistic for the other. Since many econometrics softwares provide the marginal (and sometime the joint) test statistics, the adjusted versions can be obtained effortlessly.

Significance of $RS_{\psi\phi}$ indicates some form of misspecification in the basic model with parameter ω only. However, the correct source(s) of departure can be identified only by using the adjusted statistics RS^*_ψ and RS^*_ϕ not the marginal ones (RS_ψ and RS_ϕ). This testing strategy is close to the idea of Hillier (1991) in the sense that it partitions the overall rejection region to obtain evidence about the specific direction(s) in which the basic model needs revision. And it achieves that without estimating any of the nuisance parameters. For detailed discussion, see Sen and Bera (2011).

4.2 Score Functions and Information Matrix

Recall that for the dynamic panel spatial model, the full parameter vector was $\theta = (\beta', \sigma^2_e, \gamma, \sigma^2_\mu, \rho, \delta, \lambda, \tau)'$. In context the notation of Section 4.1, $\theta = (\omega', \psi', \phi'),'$ with $\omega = (\beta', \sigma^2_e)$ and ψ and ϕ could be any combinations of the parameters under test, namely $(\gamma, \sigma^2_\mu, \rho, \delta, \lambda, \tau)$. The restricted model (under the joint null, $H^a_0 : \gamma = \sigma^2_\mu = \rho = \delta = \lambda = \tau = 0$) is simple panel model, i.e., $y_{it} = X_{it}\beta + u_{it}$, where, $u_{it} \sim IIDN(0, \sigma^2_u)$. For simplicity I assume the weight matrices W and M to be same, which is often realistic in practice. The score functions and information matrix J evaluated at the restricted MLE of θ with $\tilde{\omega} = (\tilde{\beta}, \tilde{\sigma^2_e})$ are:

$$\frac{\partial L}{\partial \beta} = 0 \tag{17}$$

$$\frac{\partial L}{\partial \sigma^2_e} = 0 \tag{18}$$

$$\frac{\partial L}{\partial \gamma} = \frac{\tilde{u}'[I_T \otimes Y_{NT-1}]}{\tilde{\sigma_e^2}} \tag{19}$$

$$\frac{\partial L}{\partial \sigma_\mu^2} = \frac{NT}{2\tilde{\sigma_e^2}}[\frac{\tilde{u}'(J_T \otimes I_N)\tilde{u}}{\tilde{u}'\tilde{u}} - 1] \tag{20}$$

$$\frac{\partial L}{\partial \rho} = \frac{NT}{2}[\frac{\tilde{u}'(G \otimes I_N)\tilde{u}}{\tilde{u}'\tilde{u}}] \tag{21}$$

$$\frac{\partial L}{\partial \delta} = \frac{\tilde{u}'[(I_T \otimes W)Y_{NT-1}]}{\tilde{\sigma_e^2}} \tag{22}$$

$$\frac{\partial L}{\partial \tau} = \frac{\tilde{u}'[(I_T \otimes W)Y_{NT}]}{\tilde{\sigma_e^2}} \tag{23}$$

$$\frac{\partial L}{\partial \lambda} = \frac{NT}{2}[\frac{\tilde{u}'(I_T \otimes (W + W'))\tilde{u}}{\tilde{u}'\tilde{u}}] \tag{24}$$

where $\tilde{u} = y - x\tilde{\beta}$ is the OLS residual vector of dimension $NT \times 1$, $\tilde{\sigma_e^2} = \frac{\tilde{u}'\tilde{u}}{NT}$ and $G = \frac{\partial V_1}{\partial \rho}|_{H_o^a}$, where G is bidiagonal matrix with bidiagonal elements all equal to one. Y_{NT} and Y_{NT-1} are vector of y and lagged values of y respectively, each of dimension $NT \times 1$.

The information matrix J, equation (11), under H_o^a is

$$J(\theta_o) = \begin{bmatrix} J_\beta & 0 & J_{\beta\gamma} & 0 & 0 & J_{\beta\delta} & J_{\beta\tau} & 0 \\ 0 & J_{\sigma_e^2} & J_{\sigma_e^2\gamma} & J_{\sigma_e^2\sigma_\mu^2} & 0 & 0 & 0 & 0 \\ J_{\gamma\beta} & J_{\gamma\sigma_e^2} & J_\gamma & J_{\gamma\sigma_\mu^2} & J_{\gamma\rho} & 0 & 0 & 0 \\ 0 & J_{\sigma_\mu^2\sigma_e^2} & J_{\sigma_\mu^2\gamma} & J_{\sigma_\mu^2} & J_{\sigma_\mu^2\rho} & 0 & 0 & 0 \\ 0 & 0 & J_{\rho\gamma} & J_{\rho\sigma_\mu^2} & J_\rho & 0 & 0 & 0 \\ J_{\delta\beta} & 0 & 0 & 0 & 0 & J_\delta & J_{\delta\tau} & J_{\delta\lambda} \\ J_{\tau\beta} & 0 & 0 & 0 & 0 & J_{\tau\delta} & J_\tau & J_{\tau\lambda} \\ 0 & 0 & 0 & 0 & 0 & J_{\lambda\delta} & J_{\lambda\tau} & J_\lambda \end{bmatrix} \tag{25}$$

where $J = E(-\frac{1}{NT}\frac{\partial^2 L}{\partial\tilde{\theta}\partial\tilde{\theta}'})$ evaluated at θ_0. The detailed derivation and expression of each of the terms of the information matrix are relegated to the appendix.

Apart from the RS statistic for full joint null hypothesis H_0^a, I propose *six* (modified) test statistics for the following hypotheses:

I) $H_o^b : \gamma = 0$ in presence of $\sigma_\mu^2, \rho, \delta, \tau, \lambda$.

II) $H_o^c : \sigma_\mu^2 = 0$ in presence of $\gamma, \rho, \delta, \tau, \lambda$.

III) $H_o^d : \rho = 0$ in presence of $\gamma, \sigma_\mu^2, \delta, \tau, \lambda$.

IV) $H_o^e : \delta = 0$ in presence of $\gamma, \sigma_\mu^2, \rho, \tau, \lambda$.

V) $H_o^f : \tau = 0$ in presence of $\gamma, \sigma_\mu^2, \rho, \delta, \lambda$.

VI) $H_o^g : \lambda = 0$ in presence of $\gamma, \sigma_\mu^2, \rho, \delta, \tau$.

Decision on testing these six hypotheses can guide a researcher to identify the correct source(s) of departure(s) from H_o^a when it is rejected. One can test various combinations of I) to VI) by testing on two/three/four parameters at a time under the null and compute additional *ninety* test statistics ($C_2^6 + C_3^6 + C_4^6 = 90$). I will demonstrate that it is unnecessary as these six (I -VI) tests are "sufficient" to detect any misspecification in the basic model. Also keeping the total number of tests to a minimum is beneficial to avoid the pre-testing problem. Since by construction the proposed tests are independent of each other, so one can easily compute the overall significance level.

Given the full model specification, it is of-course possible to derive conditional RS and LR tests, say for $\lambda = 0$ in presence of $\gamma, \sigma_\mu^2, \rho, \delta$ and τ, however that would entail the estimation of $\gamma, \sigma_\mu^2, \rho, \delta$ and τ parameters and also of λ for LR test. For the adjusted RS test these estimations are not required as it indirectly takes care of the possible presence of nuisance parameters through the Fisher-Rao score function.

Let us take the case for $H_o^g : \lambda = 0$ in presence of $\phi = (\gamma, \sigma_\mu^2, \rho, \delta, \tau)$, i.e., testing for spatial error dependence in presence of time dynamics (γ), random effect (σ_μ^2), serial correlation (ρ), space-time dynamics (δ) and spatial lag dependence (τ). For this hypothesis, the term $J_{\psi\phi.\omega}$, i.e., $J_{\lambda\phi.\omega} \neq 0$ where $\phi = (\gamma, \sigma_\mu^2, \rho, \delta, \tau)'$ and $\omega = (\beta', \sigma_e^2)'$. The term $J_{\lambda\phi.\omega}$ can be interpreted as partial covariance of scores of λ and ϕ after eliminating the linear effect of ω. Therefore, the parameter λ is not "independent" of $(\gamma, \sigma_\mu^2, \rho, \delta, \tau)'$ and vice versa. Thus, the marginal RS test statistic based on the score d_λ, i.e., RS_λ for $H_0^b : \lambda = 0$ assuming $\phi = (\gamma, \sigma_\mu^2, \rho, \delta, \tau) = (0, 0, 0, 0, 0)$ is not valid test under the presence of ϕ. In the next subsection, the proposed test statistic, RS_λ^*, that eliminates the effects of ϕ without estimating them, and is more appropriate test will be presented. I will further show, that test statistic for λ is dependent on δ and τ only, i.e., it is asymptotically independent of $(\gamma, \sigma_\mu^2, \rho)$. Thus even if one is interested to evaluate conditional RS test for H_o^g , then estimation of all the parameters are not necessary as the test statistic for λ in presence of ϕ is only affected by the presence of other spatial parameters, i.e., δ and τ and not by *panel* parameters. These type of analysis of inter-dependencies cannot be done using LR tests. The proposed adjusted tests make this possible in an elegant way. Of course, given the current computing power, it is not difficult to estimate a complex model; however, it could be sometime hard to ensure the stability of many parameter estimates. Also theoretically the stationarity regions of the parameter space have not been fully worked for the spatial model [See Elhorst 2010].

4.3 Adjusted RS tests

In this section I present the adjusted test statistics one-by-one, each of which asymptotically follows χ_1^2 distribution under the respective null hypothesis. Detailed derivation are in the appendix.

I) $H_o^b : \gamma = 0$ in presence of $\phi = (\sigma_\mu^2, \rho, \delta, \tau, \lambda)$.

To recall, here I am testing the significance of time-dynamics (γ) in presence of random effects (σ_μ^2), serial correlation (ρ), and spatial dependence (δ, τ, λ). Using the information matrix in (25), $J_{\gamma\phi.\omega}$, which can be interpreted as a covariance between parameter of interest, i.e., γ and rest of the parameters, i.e., $\phi = (\sigma_\mu^2, \rho, \delta, \tau, \lambda)$ is given by $J_{\gamma\phi.\omega} = (J_{\gamma\sigma_\mu^2.\sigma_e^2}, J_{\gamma\rho}, 0, 0, 0)$. From this we can infer:

(i) unadjusted RS test for H_o^b is not valid.

(ii) the partial covariance of d_γ and $d_{\sigma_\mu^2}$, and, d_γ and d_ρ are nonzero; while covariance with the spatial parameters (δ, τ, λ) are zero. Thus the test for γ is affected by the presence of σ_μ^2 and ρ only.

The adjusted test, RS_γ^*, takes care of the presence of σ_μ^2 and ρ using the score function of σ_μ^2 and ρ, i.e., equations (20) and (21). These scores can be viewed as "sufficient" statistics and thus can be interpreted as the indirect estimators of the respective parameters. For example, in a simple time-series model, $\hat{\rho} = \frac{\sum \tilde{u_t} \tilde{u_{t-1}}}{\sum \tilde{u_t} \tilde{u_t}'}$, and Durbin-Watson test, which is essentially a RS test, are related by: DW $\approx 2(1 - \hat{\rho})$. Here, instead of direct estimation of the nuisance parameters, ρ and σ_μ^2 , the adjusted test utilizes their respective scores, i.e.,

$$d_{\sigma_\mu^2} = \frac{\partial L}{\partial \sigma_\mu^2} = \frac{NT}{2\tilde{\sigma_e^2}}[\frac{\tilde{u}'(J_T \otimes I_N)\tilde{u}}{\tilde{u}'\tilde{u}} - 1]$$

$$d_\rho = \frac{\partial L}{\partial \rho} = \frac{NT}{2}[\frac{\tilde{u}'(G \otimes I_N)\tilde{u}}{\tilde{u}'\tilde{u}}]$$

One can note the similarity between $\hat{\rho}$ and d_ρ. The adjusted test statistic for $H_o : \gamma = 0$ is:

$$RS_\gamma^* = \frac{[d_\gamma - J_{\gamma\sigma_\mu^2.\sigma_e^2} J^{-1}_{\sigma_\mu^2.\sigma_e^2} d_{\sigma_\mu^2} - J_{\gamma\rho} J_\rho^{-1} d_\rho]^2}{[J_{\gamma.\omega} - J_{\gamma\sigma_\mu^2.\sigma_e^2} J^{-1}_{\sigma_\mu^2.\sigma_e^2} J_{\sigma_\mu^2\gamma.\sigma_e^2} - J_{\gamma\rho} J_\rho^{-1} J_{\rho\gamma}]}. \tag{26}$$

While the unadjusted counterpart is

$$RS_\gamma = \frac{d_\gamma^2}{J_{\gamma.\omega}}. \tag{27}$$

From equation (26) it is quite apparent how RS_γ^* takes care of the possible presence of nuisance parameters (σ_μ^2, ρ) using their respective scores evaluated under joint null H_o^a. It is based on the effective score $d_\gamma^* = [d_\gamma - J_{\gamma\sigma_\mu^2.\sigma_e^2} J^{-1}_{\sigma_\mu^2.\sigma_e^2} d_{\sigma_\mu^2} J_{\gamma\rho} J_\rho^{-1} d_\rho]$ which renders d_γ^* to be orthogonal to $d_{\sigma_\mu^2}$ and d_ρ. For other nuisance parameters, (δ, τ, λ), no such adjustments are necessary as it is evident from the expression of $J_{\gamma\phi.\omega}$ that $J_{\gamma(\delta\tau\lambda)} = (0, 0, 0)$, i.e., asymptotically they do not affect γ as far as testing is concerned. Thus inference on γ is affected only by the presence of panel and time-series parameters i.e., σ_μ^2 and ρ, and *not* by the presence of any spatial parameters (δ, τ, λ). This separation between time and space parameters is quite interesting, and RS_γ^* takes full advantage of it which is not possible under other test procedures.

For the following hypotheses, I mention the respective test statistics.

II) $H_o^c : \sigma_\mu^2 = 0$ in presence of $\gamma, \rho, \delta, \tau, \lambda$.

Here, I am testing for random effects (σ_μ^2), in presence of time dynamics (γ), serial correlation of errors (ρ), space-time dynamics (δ), spatial lag dependence (τ) and spatial error dependence (λ). The crucial quantity is $J_{\sigma_\mu^2 \phi.\omega} = (J_{\sigma_\mu^2 \gamma.\sigma_e^2}, J_{\sigma_\mu^2 \rho}, 0, 0, 0)$ where $\phi = (\gamma, \rho, \delta, \tau, \lambda)$

The adjusted RS test statistics is:

$$RS^*_{\sigma_\mu^2} = \frac{[d_{\sigma_\mu^2} - J_{\sigma_\mu^2 \gamma.\sigma_e^2} J^{-1}_{\gamma.\omega} d_\gamma - J_{\sigma_\mu^2 \rho} J_\rho^{-1} d_\rho]^2}{[J_{\sigma_\mu^2.\sigma_e^2} - J_{\sigma_\mu^2 \gamma.\sigma_e^2} J^{-1}_{\gamma.\omega} J_{\gamma \sigma_\mu^2.\sigma_e^2} - J_{\sigma_\mu^2 \rho} J_\rho^{-1} J_{\rho \sigma_\mu^2}]}, \tag{28}$$

and the unadjusted one

$$RS_{\sigma_\mu^2} = \frac{d^2_{\sigma_\mu^2}}{J_{\sigma_\mu^2.\sigma_e^2}} \tag{29}$$

Here σ_μ^2 is dependent only on γ and ρ, therefore $RS_{\sigma_\mu^2}$ uses the effective score $d^*_{\sigma_\mu^2} = [d_{\sigma_\mu^2} - J_{\sigma_\mu^2 \gamma.\sigma_e^2} J^{-1}_{\gamma.\omega} d_\gamma - J_{\sigma_\mu^2 \rho} J_\rho^{-1} d_\rho]$, making it orthogonal to d_γ and d_ρ.

III) $H_o^d : \rho = 0$ in presence of $\gamma, \sigma_\mu^2, \delta, \tau, \lambda$.

Here, $\phi = (\gamma, \sigma_\mu^2, \delta, \tau, \lambda)$ and $J_{\rho\phi.\omega} = (J_{\rho\gamma}, J_{\rho\sigma_\mu^2}, 0, 0, 0)$.

Thus, the adjusted test statistic is:

$$RS^*_\rho = \frac{[d_\rho - J_{\rho\gamma.\sigma_e^2} J^{-1}_{\gamma.\omega} d_\gamma - J_{\rho\sigma_\mu^2.\sigma_e^2} J^{-1}_{\sigma_\mu^2.\sigma_e^2} d_{\sigma_\mu^2}]^2}{[J_\rho - J_{\rho\gamma.\sigma_e^2} J^{-1}_{\gamma.\omega} J_{\gamma\rho.\sigma_e^2} - J_{\rho\sigma_\mu^2.\sigma_e^2} J^{-1}_{\sigma_\mu^2.\sigma_e^2} J_{\sigma_\mu^2 \rho.\sigma_e^2}]} \tag{30}$$

and while the unadjusted one is :

$$RS_\rho = \frac{d_\rho^2}{J_\rho} \tag{31}$$

It is evident from the term $J_{\rho\phi.\omega}$, that among all the nuisance parameters, serial correlation (ρ) is directly affected by presence of only time dynamics (γ) and random effects (σ_μ^2). Thus, the effective score of the test statistic [$d_\rho^* = d_\rho - J_{\rho\gamma.\sigma_e^2} J^{-1}_{\gamma.\omega} d_\gamma - J_{\rho\sigma_\mu^2.\sigma_e^2} J^{-1}_{\sigma_\mu^2.\sigma_e^2} d_{\sigma_\mu^2}$], in equation (30), clearly reveals this, i.e., d_ρ^* is made orthogonal to d_γ and $d_{\sigma_\mu^2}$.

IV) $H_o^e : \delta = 0$ in presence of $\gamma, \sigma_\mu^2, \rho, \tau, \lambda$.

Here, $\phi = (\gamma, \sigma_\mu^2, \rho, \tau, \lambda)$ and $J_{\delta\phi.\omega} = (0, 0, 0, J_{\delta\lambda.\beta}, J_{\delta\tau.\beta})$.

The adjusted RS test statistic is:

$$RS_\delta^* = \frac{[d_\delta - J_{\delta\lambda.\beta} J_{\lambda.\beta}^{-1} d_\lambda - J_{\delta\tau.\beta} J_{\tau.\beta}^{-1} d_\tau]^2}{[J_{\delta.\beta} - J_{\delta\lambda.\beta} J_{\lambda\beta}^{-1} J_{\lambda\delta.\beta} - J_{\delta\tau.\beta} J_{\tau.\beta}^{-1} J_{\tau\delta.\beta}]} \tag{32}$$

and the unadjusted one is

$$RS_\delta = \frac{d_\delta^2}{J_{\delta.\beta}} \tag{33}$$

V) $H_o^f : \tau = 0$ in presence of $\gamma, \sigma_\mu^2, \rho, \delta, \lambda$.

The set of nuisance parameters is $\phi = (\gamma, \sigma_\mu^2, \rho, \delta, \lambda)$, and $J_{\tau\phi.\omega} = (0, 0, 0, J_{\tau\delta.\beta}, J_{\tau\lambda.\beta})$.

The adjusted test statistic is:

$$RS_\tau^* = \frac{[d_\tau - J_{\tau\delta.\beta} J_{\delta.\beta}^{-1} d_\delta - J_{\tau\lambda.\beta} J_{\lambda.\beta}^{-1} d_\lambda]^2}{[J_{\tau.\beta} - J_{\tau\delta.\beta} J_{\delta.\beta}^{-1} J_{\delta\tau.\beta} - J_{\tau\lambda.\beta} J_{\lambda.\beta}^{-1} J_{\lambda\tau.\beta}]} \tag{34}$$

The unadjusted test statistic is:

$$RS_\tau = \frac{d_\tau^2}{J_{\tau.\beta}} \tag{35}$$

Similar to the other proposed test statistics, RS_τ^* also takes care of the presence of nuisance parameters through their respective scores. This is evident from the equation (34).

Lastly,
VI) $H_o^g : \lambda = 0$ in presence of $\gamma, \sigma_\mu^2, \rho, \delta, \tau$.

Here, $\phi = (\gamma, \sigma_\mu^2, \rho, \delta, \tau)$ and the partial covariance term, i.e., $J_{\lambda\phi.\omega} = (0, 0, 0, J_{\lambda\delta}, J_{\lambda\tau})$. The proposed adjusted test takes care of the interdependence of the two nuisance parameters, space-time dynamics (δ) and spatial lag dependence (τ), through their score functions. Thus, the adjusted test statistic is:

$$RS_\lambda^* = \frac{[d_\lambda - J_{\lambda\delta} J_{\delta.\beta}^{-1} d_\delta - J_{\lambda\tau} J_{\tau.\beta}^{-1} d_\tau]^2}{[J_\lambda - J_{\lambda\delta} J_{\delta.\beta}^{-1} J_{\delta\lambda} - J_{\lambda\tau.\delta} J_{\tau.\beta}^{-1} J_{\tau\lambda}]} \tag{36}$$

The unadjusted RS is :

$$RS_\lambda = \frac{d_\lambda^2}{J_\lambda} \tag{37}$$

Using these proposed adjusted test statistics, I will address the empirical question at hand, i.e., specification search for growth model. Specifically, I will start with the basic panel model and estimate it by OLS method and then will use the proposed tests to identify the specific sources of departures. Before embarking on data analysis, I discuss some further elegant and useful features of specification search.

4.4 Analysis of Misspecification

Earlier I demonstrated that the“time and panel” parameters $(\gamma, \sigma_\mu^2, \rho)$ are orthogonal to the “spatial” parameters (δ, λ, τ) in the sense of testing. Thus the joint RS statistic for $H_o^a : \gamma = \sigma_\mu^2 = \rho = \delta = \tau = \lambda = 0$, RS_J decomposes naturally into two orthogonal components:

$$RS_J = RS_{\gamma\sigma_\mu^2\rho} + RS_{\delta\tau\lambda}. \tag{38}$$

As I noted there is no further orthogonality among (within) time and panel parameters $(\gamma, \sigma_\mu^2, \rho)$, and spatial parameters (δ, τ, λ). Thus one needs to use the adjusted tests to decompose $RS_{\gamma\sigma_\mu^2\rho}$ and $RS_{\delta\tau\lambda}$ further. From the expressions of different test statistics it follows that

$$RS_{\gamma\sigma_\mu^2\rho} = RS^*_{\gamma|\sigma_\mu^2\rho} + RS_{\sigma_\mu^2\rho} = RS^*_{\gamma|\sigma_\mu^2\rho} + RS^*_{\sigma_\mu^2|\rho} + RS_\rho = RS^*_{\gamma|\sigma_\mu^2\rho} + RS^*_{\rho|\sigma_\mu^2} + RS_{\sigma_\mu^2} \tag{39}$$

where $RS^*_{\gamma|\sigma_\mu^2\rho}$ is the adjusted test derived in equation (26), $RS^*_{\sigma_\mu^2|\rho}(RS^*_{\rho|\sigma_\mu^2})$ is the adjusted test statistics for σ_μ^2 (ρ) after taking care of the parameter ρ (σ_μ^2). Moreover, the analytical form of $RS^*_{\sigma_\mu^2|\rho}$ and $RS^*_{\rho|\sigma_\mu^2}$ are same as derived in Sen and Bera (2011) under static panel spatial model framework.

Alternatively, one can also write:

$$RS_{\gamma\sigma_\mu^2\rho} = RS^*_{\sigma_\mu^2|\gamma\rho} + RS_{\gamma\rho} = RS^*_{\sigma_\mu^2|\gamma\rho} + RS^*_{\gamma|\rho} + RS_\rho = RS^*_{\sigma_\mu^2|\gamma\rho} + RS\rho|\gamma^* + RS_\gamma \tag{40}$$

or,

$$RS_{\gamma\sigma_\mu^2\rho} = RS^*_{\rho|\gamma\sigma_\mu^2} + RS_{\gamma\sigma_\mu^2} = RS^*_{\rho|\gamma\sigma_\mu^2} + RS^*_{\gamma|\sigma_\mu^2} + RS_{\sigma_\mu^2} = RS^*_{\rho|\gamma\sigma_\mu^2} + RS^*_{\sigma_\mu^2|\gamma} + RS_\gamma \tag{41}$$

Most computer software reports joint and unadjusted (one-directional) RS tests. Above decomposition suggest that one can obtain all the adjusted test without any extra computation. Similar decomposition will hold for the adjusted RS test statistics for the spatial parameters $(\delta, \tau$ and $\lambda)$.

$$RS_{\delta\tau\lambda} = RS^*_{\delta|\tau\lambda} + RS_{\tau\lambda} = RS^*_{\delta|\tau\lambda} + RS^*_{\tau|\lambda} + RS_\lambda = RS^*_{\delta|\tau\lambda} + RS^*_{\lambda|\tau} + RS_\tau \tag{42}$$

Therefore, the proposed adjusted test statistics aid the researcher in model specification with minimum estimation, and its elegant additive property give the researcher a wide flexibility in selecting a model framework.

5 Specification Search for Growth Model

5.1 Data

The data is from Penn World Tables (PWT, version 6.1), which contain information on real income, investment and population (among many other variables) for a large number of countries. In this paper, I use a sample of 91 countries over the period of 1961 - 1995. These countries are those of MRW (1992) non-oil sample which has been used extensively by other researchers for empirical work on growth convergence.

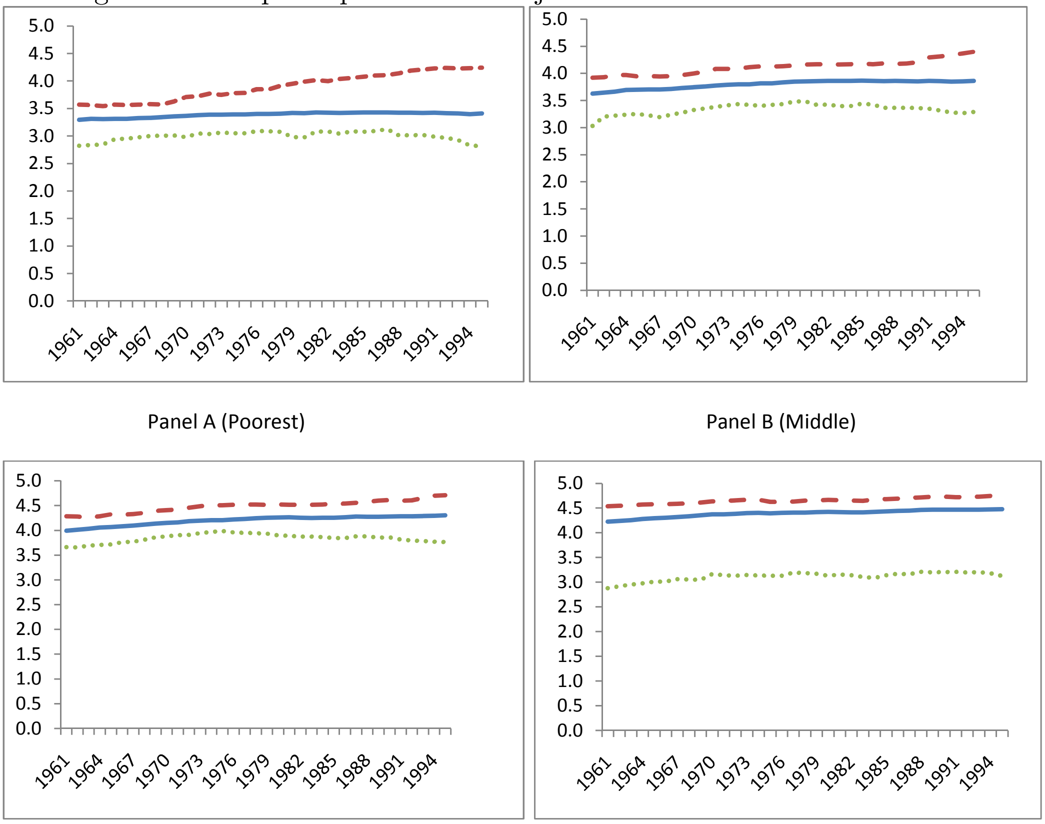
The dependent variable is real income per worker is measured by real GDP computed by chain method, divided by number of workers. I computed the number of workers following Caselli (2005): $RGDPCH \times POP/RGDPW$, where RGDPCH is real GDP per capita computed by chain method, RGDPW is real GDP per worker and POP is population. The independent variables are same as in MRW (1992). They are n, which measures the average growth of the working-age population (ages 15 to 64), the savings rate s is measured as the average share of gross investment in GDP.

There are many ways one can specify the weight matrix W, for example, geographic distance, k-neighborhood matrix, contiguous neighborhood matrix, economic distance, etc. (See Conley and Topa (2002), Eaton and Kortun (1996), Klenow and Rodriguez-Clare (2005) and Ertur and Koch (2007).) I consider three different specification of weight matrix, mainly to check the sensitivity of the test result to different definition of spatial connectedness. Here, W_1, W_2 and W_3, where W_1 are defined. The elements of $W_1 are\ w1_{ij} = \frac{d_{ij}^{-2}}{\sum_j d_{ij}^{-2}}$, such that $d_{ii} = 0$ and d_{ij} is the euclidean distance between country's capital. Other two matrices, W_2 and W_3, are based on k-nearest neighbors, with $k = 8$ and 20 respectively, nearness being measured in terms of the geographic distance.

5.2 Specification Search

First I present some basic features of the income distribution of the 91 non-oil countries over the 35- year period 1961 - 1995, in Figure 1 for four groupings of cross sectional averages of per capita real income, where the groups are selected based on initial income of these countries in 1961. [The details of each group is provided in the appendix]. Averages, maximum and minimum are shown across the four panels. Panel A, B, C, D are respectively, for the poorest, middle, rich and richest income groups. A, B and C each is based on 24 countries, and Panel D represents 19 countries. The trajectories in Figure 1 provide some idea of the variability in the actual growth trajectories over time within these groupings. It also indicate that some members of each group have substantial prospects of moving into higher income groups over the 35 year period. However,

Figure 1: Real per-capita income trajectories of 91 countries: 1961 - 1995



Panel A (Poorest) Panel B (Middle)

Panel C (Rich) Panel D (Richest)

Average (blue ——), Minimum (green......) and Maximum (red - - - -)

assuming homogeneity of technological progress and speed of convergence (which is usually assumed for β- convergence in cross-sectional studies) rules out this possibility. It is only when one allows for heterogeneity in technological progress over time and over cross-sectional units (while at the same time requiring that the growth rate of technological progress converge to a common constant over time to ensure convergence), then the realistic patterns as shown in Figure 1 can emerge (for detail, see Philips and Sul (2003)). Using the proposed specification tests our objective is now to identify a model that captures such essential features of the data.

Table 1 reports the joint RS test RS_J for the null $H_o^a : \gamma = \sigma_\mu^2 = \rho = \delta = \tau = \lambda = 0$, joint test for time dynamics (γ), random effect (σ_μ^2) and serial correlation (ρ) $RS_{\gamma\sigma_\mu^2\rho}$, joint RS test for space recursive (δ), spatial lag (τ) and spatial error lag (λ) $RS_{\delta\tau\lambda}$ and Table 2 reports the unadjusted single-directional RS tests for each of the six parameters, and the proposed adjusted RS test statistics (noted by '*') for all the parameters. Except $RS_J, RS_{\gamma\sigma_\mu^2\rho}$ and $RS_{\delta\tau\lambda}$, each of the test statistics follow χ_1^2 distribution asymptotically. $RS_J \sim \chi_6^2$, and $RS_{\gamma\sigma_\mu^2\rho}$ and $RS_{\delta\tau\lambda} \sim \chi_3^2$. Each of these tables report the test statistics for same model under three different specification of the weight matrix W. Recall, W_1 uses the geographical distance between the capital of the countries, W_2 and W_3 use eight and twenty nearest neighbors, respectively.

Table 1: Specification Search using Full Sample of 91 countries over 1961 - 1995.

Test Statistics	Specification with W_1	Specification with W_2	Specification with W_3
RS_J	231.91	177.16	217.16
$RS_{\gamma\sigma_\mu^2\rho}$	188.19	141.19	179.19
$RS_{\delta\tau\lambda}$	43.72	35.97	37.97

Table 2: Specification Search using Full Sample of 91 countries over 1961 - 1995.

Parameters	RS_{W_1}	$RS^*_{W_1}$	RS_{W_2}	$RS^*_{W_2}$	RS_{W_3}	$RS^*_{W_3}$
Time-Dynamics - γ	81.12***	48.11***	60.09***	37.51***	79.51***	43.75***
Heterogeneity - σ_μ^2	92.45***	59.17***	79.12***	43.68***	89.03***	54.84***
Serial Correlation - ρ	32.31***	15.79***	24.69***	9.94***	27.23***	12.55***
Space-Time Dyn - δ	8.51***	**1.11**	6.42**	**0.19**	7.50***	**0.12**
Spatial Lag - τ	7.96***	**2.42**	5.91**	**1.74**	6.96***	**1.82**
Spatial Error - λ	41.92***	40.68***	33.31***	31.17***	36.14***	35.57***

Note: * indicates significant at 10%, ** indicates significant at 5% and *** indicates significant at 1%.

No matter which W is chosen, RS_J is always highly significant at any significance level. Given the orthogonality between spatial and panel-time parameters, as given in the additivity result in equation (38), we can conduct the joint tests $RS_{\gamma\sigma_\mu^2\rho}$ and $RS_{\delta\tau\lambda}$. These joint tests are also highly significant after comparing with χ_3^2 critical points at any significance level; however they are not informative about the specific direction(s) of the misspecification(s). Thus based on the results for the unadjusted tests, it would appear that the features like time dynamics, serial correlation of error, random effects, spatial dependencies are the features of this dataset and therefore should be added to the basic model (joint null). However, as we discussed earlier, the inference based on the unadjusted tests can be highly misleading as they fail to take into account the possible presence of other parameters and their interdependencies.

Significance of each parameter can only be evaluated correctly by considering the modified tests. Out of all the six adjusted test statistics, only four, namely, $RS^*_\gamma, RS^*_{\sigma^2_\mu}, RS^*_\rho$ and RS^*_λ are significant, irrespective of the choice of weight matrix W. It is interesting to note the difference in values of the test statistics; the adjusted test statistics are much lower than their unadjusted counterparts. The striking differences in values can be noted for RS_δ, RS_τ with their adjusted counterparts, i.e., RS^*_δ and RS^*_τ respectively. The value of RS^*_δ falls below the critical point of χ^2_1 at any significance level, after it takes into account the possible presence of τ and λ. Similarly spatial lag dependence (τ) parameter loses its significance as the test statistic drops from 7.96 to 2.42 after adjustment. Viewing the test statistics as a measure of degree of misspecification, we find that out of all the departures from the joint null, most of the misspecification is attributed to σ^2_μ ($\frac{59.17}{231.91} = 26\%$) and time dynamics γ ($\frac{48.11}{231.91} = 21\%$) followed by spatial error dependence (which captures the indirect cross-sectional dependence among these countries). This holds true with W_3; for W_2, however, when the weight matrix is relatively sparse. Thus for W_2, the misspecification due to spatial dependence is relatively low. This may be due to the fact that W_2 is sparse and thus dilute the degree of spatial dependencies. However, no matter what form of W is chosen, misspecification due to time dynamics, serial correlation and random effects are always strong. Thus, when the full sample of annual data is used for 91 non-oil countries for the period 1961 - 1995, the relevant growth regression [2] would be

$$y_{it} = \gamma y_{it-1} + \beta_1 x_{it} + \mu_i + \epsilon_{it} \tag{43}$$

$$\epsilon_{it} = \lambda \sum_{j \neq i} w_{ij}\epsilon_{jt} + v_{it} \tag{44}$$

$$v_{it} = \rho v_{it-1} + e_{it} \tag{45}$$

Note that this specification, interestingly, also supports kind of trajectories Figure 1 demonstrates as it can take into account the heterogeneity in technological progress across countries and across time, captured through the spatial dependence (which can account for the technology

[2]Concerns may be raised regarding the estimation of such model specially when the time dynamics is present along with the random effect, as it is generally believed that the inclusion of lagged dependent variables in a panel model necessarily renders random-effect estimators inconsistent. However, it has been shown in Ashley (2010) that if the variables X are strictly exogenous then the lagged value of the quasi-differenced dependent variable is uncorrelated with the quasi-differenced model error, and thus usual random effect estimator would provide consistent estimators of the parameters.

transfer between countries) and the serial correlation (which captures the differences across time).[3].

MRW (1992) considers a cross-sectional regression model for the non-oil sample with s and n as explanatory variables and finds the rate of conditional convergence be to be very low, 0.00606(0.001) (speed of convergence $= -(1-\hat{\alpha})(n+g+\xi)$, where $\hat{\alpha}$ is the estimated share of physical capital, n is working population, g is growth rate of the country, ξ is depreciation of capital), implying a half-life of 114 years, which is indeed very long. Islam (1995) uses fixed effect dynamic panel data model and allows for the unobserved technological diffusion through the fixed effects term, and estimates the rate of conditional convergence to be 0.0434 (0.007) (speed of convergence measured as $\frac{1}{\Delta}\ln\hat{\gamma}$, where Δ is the time difference between two consecutive periods and $\hat{\gamma}$ is the estimate of time dynamics parameter in a fixed effect dynamic panel data model) . It should be noted here, that Islam (1995) used minimum distance (MD) estimator to estimate his model using similar data as MRW(1992). According to Islam (1995), the panel estimate of the convergence rate increases 7.2 times (relative to its OLS estimates that ignores technological differences, as in MRW (1992)) in the non-oil sample, thus concluding that for these countries the half-life is 16 years approximately. Lee et al. (1997) estimate the rate of convergence to be 0.1845 for the same sample of countries by allowing the growth rate g to differ across countries and also for possible serial correlation of error. Ertur and Koch (2007) consider the growth convergence model allowing for regional knowledge and technology spillover effects through spatial dependence, and estimate the rate of convergence to be 0.012 (0.00), with half life around 59 years (speed of convergence $= \frac{\sum_{j=1}^{N} u_{ij}\frac{1}{\Phi_j}(n_j+g+\xi)}{\sum_{j=1}^{N} u_{ij}\frac{1}{\Phi_j}} - \sum_{j=1}^{N} u_{ij}\frac{1}{\Theta_j}(n_j+g+\xi)$, where u_{ij} is a function of estimates of capital share, spatial dependence and elements of W matrix, Φ_j and Θ_j are the rate of convergence of capital and income, respectively, to the steady state, of country j. For details see, Ertur and Koch (2007).) Thus, the speed of convergence and the implied half-life clearly depends on the model framework. For same dataset, under different frameworks, researchers got widely varying rates of convergence and widely varying half-life estimates (number of years needed forconditional convergence) corresponding to each of the respective convergence estimates. This is a surprising fact and hasn't been considered so far by any research papers. So it is obvious that these models could not capture all the salient feature of the underlying data and thats why the estimate of growth convergence from these models can result in potentially misleading policy implications. The proposed tests can aid the researcher in tackling this difficult task- i.e., to understand the DGP with minimum estimation a priori. As I discussed, given the annual data for the sample of 91 countries over the period 1961 - 1995, the most appropriate model specification is given by equations (43) - (45).

In Table 3 and 4, I have considered specification search for growth models under different time frames. I consider the test statistics when the time span is five - year intervals. Thus considering the period 1961 - 1995, I have seven data (time) points for each country: 1965, 1970, 1975, 1980, 1985,

[3]Only assumption required for convergence here is that over time (technically $t \to \infty$) the technological differences between countries i and j should go to zero.

1990, 1995. The variables are averages over five -year time intervals. Islam (1995) pointed out that "yearly time span are too short to be appropriate for growth convergence. Short-term disturbances may loom large in such brief time spans." Following his work, many researchers considered this data setup so that the growth convergence estimates are less influenced by business cycle fluctuations and less likely to be serially correlated than they would be in a yearly data setup. I also divide the data in two subsamples to investigate if the the model specification search is robust. Therefore, RS^*_{S1}, RS_{S1}, RS^*_{S2} and RS_{S2} are respectively the adjusted and unadjusted test statistics for the annual data for 91 countries for the subsamples 1961 - 1980 (referred as S1) and 1981 - 1995 (referred as S2) respectively.

Table 3: Specification search using different time specification.

Test Statistics	5-year time interval	Subsample: 1961 - 1980	Subsample: 1981 - 1995
RS_J	87.41	170.82	154.86
$RS_{\gamma\sigma^2_\mu\rho}$	70.13	145.14	131.47
$RS_{\delta\tau\lambda}$	17.28	25.68	23.39

Table 4: Specification search using different time specification.

Parameters	$RS_{5-years}$	$RS^*_{5-years}$	RS_{S1}	RS^*_{S1}	RS_{S2}	RS^*_{S2}
Time-Dynamics - γ	18.50***	10.96***	72.29***	37.11***	69.51***	33.75***
Heterogeneity - σ^2_μ	25.27***	14.07***	89.31***	41.11***	80.23***	39.17***
Serial Correlation - ρ	14.01***	**2.80**	21.69***	6.94***	24.23***	7.55***
Space-Time Dyn - δ	8.05***	**1.31**	5.33**	**1.56**	7.68***	**1.87**
Spatial Lag - τ	**1.01**	**0.009**	4.79**	**1.19**	6.45***	**1.43**
Spatial Error - λ	15.16***	14.22**	22.11***	20.17***	21.14***	19.81***

Note: * indicates significant at 10%, ** indicates significant at 5% and *** indicates significant at 1%.
S1: Subsample: 1961 - 1980, S2: Subsample: 1980 - 1995.

It is evident from Table 3 that all the joint tests are significant irrespective of the time span of the data. From column 2 of Table 4, it is evident that only the heterogeneity (σ^2_μ) and time dynamics (γ) are most prominent features when 5 year time interval is chosen. RS^*_λ is significant at 5% level. Interestingly although the unadjusted test for ρ is significant, but adjusted one is no longer so. Thus when the data setup is based on 5-year time interval, the relevant growth equation is very similar to Islam (1995), i.e., dynamic panel data framework augmented for cross-sectional dependence also. Indeed 5 -year time span removes effect of serial correlation of errors.

Column 4 - Column 7 of Table 4 indicate that the relevant feature of growth regression are similar to equation (43) - (45). Although the relative values of the test statistics are different, but the inference remains same. This supports the robustness of the results using the proposed test statistics.

To summarize, in this section I use my proposed test statistics from Section 4, for the proper specification search of growth model. As explained in Section 4.1, I need to estimate *only* the simple panel model in order to apply the proposed tests. Thus, no complex estimation is necessary. I show how one can unravel the salient features of the underlying DGP using the proposed tests. In particular, I show that for the given dataset of 91 non-oil samples from Penn World table, the most relevant features are heterogeneity, time dynamics and indirect cross-sectional dependence. Thus a researcher analyzing the growth behavior of these 91 countries should take care of these departures in his/her model; otherwise the model would be misspecified which would lead to wrong policy implication. For example, as I have shown here, if one assumes a cross-sectional regression model, then the growth convergence rate is very low, implying the half-life to be 114 years. Again assuming a fixed effect dynamic panel model will yield a much higher rate of convergence for the same dataset, implying half-life to be as short as 16 years. It is evident that the convergence rate of income vary wildly, even when same dataset is used. Thus one should consider a proper specification search before directly going into model implication and policy analysis. Tables 1 and 2 illustrate this important fact and Tables 3 and 4 demonstrate the robustness of the proposed test result.

In the next section I demonstrate that though the suggested tests are valid only for large samples and local misspecification, they perform quite well in finite samples.

6 Monte Carlo Results

The proposed tests are valid only asymptotically. As in our empirical application, in the real world data will be limited. Therefore, we need to evaluate the performance of the tests under a finite sample scenario. The data for Monte Carlo study were generated based on the model (1) - (4).

$$y_{it} = \gamma y_{it-1} + \tau \sum_{j=1}^{N} m_{ij} y_{jt} + \delta \sum_{j=1}^{N} m_{ij} y_{jt-1} + X_{it}\beta + u_{it} \tag{46}$$

$$u_{it} = \mu_i + \epsilon_{it} \tag{47}$$

$$\epsilon_{it} = \lambda \sum_{j=1}^{N} w_{ij} \epsilon_{jt} + v_{it} \tag{48}$$

$$v_{it} = \rho v_{it-1} + e_{it}, \text{where } e_{it} \sim IIDN(0, \sigma_e^2) \tag{49}$$

We set $\alpha = 5$ and $\beta = 0.5$. The independent variable X_{it} is generated using:

$$X_{it} = 0.4 X_{it-1} + \varphi_{it} \tag{50}$$

where $\varphi \sim Unif[-0.5, 0.5]$ and $X_{i0} = 5+10\varphi_{i0}$. For weight matrix W, I consider rook design. I fixed $\sigma_\mu^2 + \sigma_e^2 = 20$ and let $\eta = \frac{\sigma_\mu^2}{\sigma_\mu^2+\sigma_e^2}$. Values of all the six parameters $\gamma, \sigma_\mu^2, \rho, \delta, \tau$ and λ are varied over a range from 0 to 0.5 . I have considered two different pairs for (N, T) namely $(25, 10), (49, 20)$. For lack of space I report the results for $(25, 10)$. The results for $(49, 20)$ are quite comparable to the reported ones and are available on request. Each Monte Carlo experiment is consist of generating 1000 samples for each different parameter settings. Thus the maximum standard error of the estimates of the size and power would be $\sqrt{\frac{(0.5(1-0.5)}{1000}} = 0.015$. The parameters were estimated using OLS, and fifteen test statistics, namely $RS_J, RS_{\gamma\sigma_\mu^2\rho}, RS_{\delta\tau\lambda}, RS_\gamma^*, RS_\gamma, RS_{\sigma_\mu^2}^*, RS_{\sigma_\mu^2}, RS_\rho^*, RS_\rho, RS_\delta^*, RS_\delta, RS_\tau^*, RS_\tau, RS_\lambda^*$ and RS_λ were computed. As discussed earlier, in practice it is not necessary to compute all these statistics ; I do it here only for comparative evaluation. The results are based on the nominal size of 0.05.

As noted in Section 4, the parameters $(\gamma, \sigma_\mu^2, \rho)$ are orthogonal to (δ, τ, λ) as far as testing are concerned. So I report the key results in two tables. Table 5 reports the size and power of $RS_\gamma^*, RS_\gamma, RS_{\sigma_\mu^2}^*, RS_{\sigma_\mu^2}, RS_\rho^*, RS_\rho$ and Table 6 reports the size and power of $RS_\delta^*, RS_\delta, RS_\tau^*, RS_\tau, RS_\lambda^*$ and RS_λ. Results $RS_J, RS_{\gamma\sigma_\mu^2\rho}$ and $RS_{\delta\tau\lambda}$ are not reported for lack of space. However, each of them achieves nominal size under joint null, and expected power properties.

In Table 5, I vary the parameters (γ, η, ρ) from 0 to 0.5, one and two at a time, keeping the spatial parameters zero. RS_γ^* is size robust while RS_γ performs badly when $\gamma = 0$ **and** when either or both $\eta \neq 0$ and $\rho \neq 0$. For example, when $\gamma = 0, \eta = 0, \rho = 0.3$, rejection probability of RS_γ^* is 0.048 and that for RS_γ is 0.888. Similarly, when $\gamma = 0.3, \eta = 0, \rho = 0$ then rejection probability of $RS_{\sigma_\mu^2}^*$ is 0.054 and that of $RS_{\sigma_\mu^2}$ is 0.231. When $\gamma = 0.4, \eta = 0, \rho = 0$ the rejection probability of RS_ρ^* is 0.040 and that of RS_ρ is 0.166. Thus as expected by construction, the adjusted test statistics are size-robust under (local and in some cases even global) misspecification while their unadjusted counter parts are not. However, there is slight loss in power for these adjusted test statistics compared to the unadjusted ones, when adjustments are made even when there is no misspecification. This loss in power however reduces as the parameter values deviates further from the null. As discussed in Sen and Bera (2011), this loss in power can be regarded as the premium one pays for the validity of the adjusted test under local misspecification, i.e., the cost of robustness.

In Table 6, I vary the spatial parameters (δ, τ, λ) from 0 to 0.5, one and two at a time, keeping other parameters at zero. It is evident from Table 6 that RS_δ^*, RS_τ^* and RS_λ^* are more size-robust than RS_δ, RS_τ and RS_λ respectively. For instance, when $\delta = 0, \tau = 0.4, \lambda = 0.4$, the rejection probability of RS_δ^* is 0.049, whereas for RS_δ it is 0.997. Again when $\delta = 0.4, \tau = 0, \lambda = 0$, the rejection probability for RS_λ^* is 0.038 and that for RS_λ is 0.99. Further Monte Carlo results on $RS_\gamma^*, RS_\gamma, RS_{\sigma_\mu^2}^*, RS_{\sigma_\mu^2}, RS_\rho^*, RS_\rho, RS_\delta^*, RS_\delta, RS_\tau^*, RS_\tau, RS_\lambda^*$ and RS_λ are reported in the Appendix.

Table 5: Estimated Rejection Probabilities with $\delta = \tau = \lambda = 0$. Sample size: $N = 25, T = 10$

γ	η	ρ	RS^*_γ	RS_γ	$RS^*_{\sigma^2_\mu}$	$RS_{\sigma^2_\mu}$	RS^*_ρ	RS_ρ
0	0	0	**0.055**	**0.057**	**0.058**	**0.051**	**0.058**	**0.054**
0.1	0	0	0.058	0.497	0.069	0.091	0.115	0.198
0.2	0	0	0.060	0.769	0.054	0.107	0.105	0.291
0.3	0	0	0.048	0.888	0.035	0.131	0.267	0.316
0.4	0	0	0.049	1.000	0.066	0.108	0.277	0.460
0.5	0	0	0.044	1.000	0.041	0.344	0.314	0.698
0	0.4	0	0.054	0.682	0.119	0.099	0.034	0.110
0.1	0.4	0	0.068	0.798	0.155	0.112	0.296	0.332
0.2	0.4	0	0.084	0.894	0.151	0.202	0.287	0.413
0.3	0.4	0	0.078	0.959	0.238	0.346	0.375	0.534
0.4	0.4	0	0.081	1.000	0.324	0.425	0.451	0.611
0.5	0.4	0	0.091	1.000	0.211	0.492	0.524	0.723
0	0.1	0	0.065	0.753	0.161	0.212	0.045	0.111
0	0.2	0	0.065	0.861	0.166	0.202	0.032	0.203
0	0.3	0	0.072	0.952	0.191	0.346	0.035	0.298
0	0.4	0	0.097	0.955	0.206	0.425	0.044	0.229
0	0.5	0	0.069	0.965	0.263	0.492	0.054	0.398
0.4	0	0	0.014	1.000	0.043	0.109	0.230	0.157
0.4	0.1	0	0.015	1.000	0.113	0.129	0.239	0.264
0.4	0.2	0	0.019	1.000	0.217	0.258	0.329	0.319
0.4	0.3	0	0.029	1.000	0.217	0.254	0.354	0.401
0.4	0.4	0	0.029	1.000	0.323	0.432	0.448	0.503
0.4	0.5	0	0.045	1.000	0.343	0.521	0.570	0.512
0	0	0.1	0.153	0.745	0.078	0.071	0.049	0.109
0	0	0.2	0.127	0.865	0.067	0.207	0.044	0.111
0	0	0.3	0.231	1.000	0.054	0.231	0.046	0.209
0	0	0.4	0.226	1.000	0.045	0.208	0.040	0.166
0	0	0.5	0.334	1.000	0.066	0.344	0.043	0.254
0.4	0	0	0.044	1.000	0.053	0.111	0.221	0.188
0.4	0	0.1	0.302	1.000	0.091	0.201	0.210	0.582
0.4	0	0.2	0.401	1.000	0.061	0.294	0.311	0.889
0.4	0	0.3	0.504	1.000	0.051	0.363	0.333	0.965
0.4	0	0.4	0.534	1.000	0.076	0.388	0.355	0.994
0.4	0	0.5	0.621	1.000	0.056	0.424	0.368	0.996

Table 6: Estimated Rejection Probabilities with $\gamma = \sigma_\mu^2 = \rho = 0$. Sample size: $N = 25, T = 10$

δ	τ	λ	RS_δ^*	RS_δ	RS_τ^*	RS_τ	RS_λ^*	RS_λ
0	0	0	**0.049**	**0.056**	**0.053**	**0.054**	**0.047**	**0.061**
0	0	0.1	0.047	0.694	0.161	0.783	0.052	0.953
0	0	0.2	0.045	0.786	0.133	0.950	0.157	0.995
0	0	0.3	0.051	0.964	0.255	0.992	0.372	1.000
0	0	0.4	0.061	0.933	0.235	1.000	0.635	1.000
0	0	0.5	0.039	0.863	0.356	1.000	0.861	1.000
0	0.4	0	0.051	1.000	0.965	1.000	0.031	1.000
0	0.4	0.1	0.038	1.000	0.990	1.000	0.044	1.000
0	0.4	0.2	0.039	1.000	0.995	1.000	0.128	1.000
0	0.4	0.3	0.047	0.992	0.993	1.000	0.229	1.000
0	0.4	0.4	0.057	0.980	0.982	1.000	0.464	1.000
0	0.4	0.5	0.060	0.943	0.939	1.000	0.780	1.000
0	0.1	0	0.037	0.782	0.284	0.893	0.058	0.759
0	0.2	0	0.031	0.876	0.867	0.992	0.070	0.996
0	0.3	0	0.051	0.934	0.934	1.000	0.067	0.999
0	0.4	0	0.051	1.000	0.961	1.000	0.065	1.000
0	0.5	0	0.041	1.000	0.979	1.000	0.078	1.000
0	0.1	0.4	0.048	0.951	0.950	1.000	0.648	1.000
0	0.2	0.4	0.053	0.968	0.964	1.000	0.640	1.000
0	0.3	0.4	0.037	0.974	0.971	1.000	0.487	1.000
0	0.4	0.4	0.049	0.994	0.985	1.000	0.479	1.000
0	0.5	0.4	0.056	0.986	0.990	1.000	0.468	1.000
0.1	0	0	0.238	0.798	0.114	0.712	0.055	0.789
0.2	0	0	0.333	0.894	0.267	0.871	0.068	0.849
0.3	0	0	0.357	0.976	0.334	0.967	0.054	0.977
0.4	0	0	0.459	1.000	0.261	0.989	0.038	0.990
0.5	0	0	0.600	1.000	0.379	0.996	0.090	0.993
0	0	0.4	0.049	0.947	0.143	0.999	0.674	1.000
0.1	0	0.4	0.368	0.991	0.189	1.000	0.445	0.999
0.2	0	0.4	0.410	1.000	0.292	1.000	0.301	1.000
0.3	0	0.4	0.505	1.000	0.395	1.000	0.295	1.000
0.4	0	0.4	0.601	1.000	0.391	1.000	0.376	1.000
0.5	0	0.4	0.632	1.000	0.492	1.000	0.448	1.000

7 Conclusion

The growth convergence debate has always occupied a central stage in economics. This is mainly because of the existence of the variety of issues regarding such models, like different forms of convergence, estimation techniques, data, variables, sample and so on. In this paper I address one specific concern, i.e., what is the most appropriate model given the data. Thus the contributions of this paper are twofold. *Firstly*, this paper develops adjusted RS test statistics, which are robust under local misspecification in a dynamic panel model allowing for cross-sectional dependence. *Secondly*, using the proposed tests I address the issue of growth model specification objectively.

To achieve these objectives, robust RS tests for time dynamics, random effect, serial correlation of errors, space-time dynamics and spatial dependence are proposed using Bera and Yoon (1993) test principle. These six adjusted tests are robust under "all" possible misspecification. This robustness is achieved without any estimation of the nuisance parameters. For example, the proposed adjusted RS test for time dynamics is made robust to the presence of random effect, serial correlation of errors, space-time dynamics and spatial dependence. I take care of these possible presence of nuisance parameters using their respective Fisher-Rao score functions. Thus, there is no need to estimate the nuisance parameters as usually it is done for conditional LM and LR tests. The proposed (robust) tests are simple to compute and interpret as they are essentially based only on OLS residuals and score functions. In addition, due to an attractive additive property, the robust tests require very little extra computation. Thus one can compute these robust tests for each parameter from the standard RS tests (joint and marginal). Due to this simplicity in terms of computation, the researchers can identify specific direction(s) to reformulate the basic growth model quite easily.

In the empirical application, using these tests, I find that most of the misspecification is attributed to heterogeneity (random effects), dynamic time effects and indirect cross-sectional dependence, irrespective of the specification of weight matrix and time span of the sample. In addition, I demonstrate how the exact nature of dependencies changes the growth model specification for different time framework. Different researchers have derived *widely different convergence rates* for the same dataset, as they considered either only cross-sectional, spatial, panel, or dynamic panel models. It is quite possible that those models cannot capture all the salient feature of the data. Using a model framework which combines all these piece-wise models considered so far in the literature, I conduct the growth model specification search using the proposed test statistics developed for the dynamic panel model with cross-sectional dependence. One should note that the proposed tests are general and can be used for many other specification search of econometric models, for example, hedonic price models, unemployment models. Lastly, through simulation study I demonstrate that the proposed tests, are not only theoretically and asymptotically valid, but can also be used in finite samples exercises where availability of data is often limited.

References

Anselin, L., A.K. Bera, R. Florax and M.J. Yoon (1996), 'Simple diagnostic tests for spatial dependence', *Regional science and urban economics* **26**(1), 77–104.

Ashley, R. (2010), 'A reconsideration of consistent estimation of a dynamic panel data model in the random effects (error components) framework', Working Paper.

Barro, R.J. and X. Sala-i Martin (1992), 'Convergence', *Journal of political Economy* pp. 223–251.

Bera, A., Y. Bilias and M. Yoon (2007), 'Adjustments of raoâĂŹs score test for distributional and local parametric misspecifications', *University of Illinois at Urbana-Champaign* .

Bera, A.K., G. Montes-Rojas and W. Sosa-Escudero (2009), 'Testing under local misspecification and artificial regressions', *Economics Letters* **104**(2), 66–68.

Bera, A.K. and M.J. Yoon (1993), 'Specification testing with locally misspecified alternatives', *Econometric theory* **9**(04), 649–658.

Bera, A.K., W. Sosa-Escudero and M. Yoon (2001), 'Tests for the error component model in the presence of local misspecification', *Journal of Econometrics* **101**(1), 1–23.

Bernard, A.B. and S.N. Durlauf (1995), 'Convergence in international output', *Journal of Applied Econometrics* **10**(2), 97–108.

Binder, M. and M.H. Pesaran (1999), 'Stochastic growth models and their econometric implications', *Journal of Economic Growth* **4**(2), 139–183.

Caselli, F. (2005), 'Accounting for cross-country income differences', *Handbook of economic growth* **1**, 679–741.

Cass, D. (1965), 'Optimum growth in an aggregative model of capital accumulation', *The Review of Economic Studies* pp. 233–240.

Conley, T.G. and E. Ligon (2002), 'Economic distance and cross-country spillovers', *Journal of Economic Growth* **7**(2), 157–187.

Conley, T.G. and G. Topa (2002), 'Socio-economic distance and spatial patterns in unemployment', *Journal of Applied Econometrics* **17**(4), 303–327.

Davidson, R. and J.G. MacKinnon (1987), 'Implicit alternatives and the local power of test statistics', *Econometrica: Journal of the Econometric Society* pp. 1305–1329.

Durlauf, S. and D. Quah (1999), *The new empirics of economic growth, in (J.Taylor and M.Woodford, eds.), Handbook of Macroeconomics*, Amsterdam: North-Holland.

Durlauf, S.N., P.A. Johnson and J.R.W. Temple (2005), 'Growth econometrics', *Handbook of economic growth* **1**, 555–677.

Eaton, J. and S. Kortum (1996), 'Trade in ideas patenting and productivity in the oecd', *Journal of international Economics* **40**(3-4), 251–278.

Elhorst, J.P. (2005), 'Unconditional Maximum Likelihood Estimation of Linear and Log-Linear Dynamic Models for Spatial Panels', *Geographical Analysis* **37**(1), 85–106.

Elhorst, J.P. (2010), 'Spatial panel data models', *Handbook of applied spatial analysis* pp. 377–407.

Ertur, C., J. Le Gallo and C. Baumont (2006), 'The european regional convergence process, 1980-1995: Do spatial regimes and spatial dependence matter?', *International Regional Science Review* **29**(1), 3.

Ertur, C. and W. Koch (2007), 'Growth, technological interdependence and spatial externalities: theory and evidence', *Journal of Applied Econometrics* **22**(6), 1033–1062.

Evans, P. and G. Karras (1996*a*), 'Convergence revisited', *Journal of Monetary Economics* **37**(2), 249–265.

Evans, P. and G. Karras (1996*b*), 'Do economies converge? evidence from a panel of us states', *The review of Economics and Statistics* pp. 384–388.

Hillier, G.H. (1991), 'On multiple diagnostic procedures for the linear model', *Journal of econometrics* **47**(1), 47–66.

Holtz-Eakin, D. (1992), 'Solow and the states: Capital accumulation, productivity and economic growth'.

Islam, N. (1995), 'Growth empirics: a panel data approach', *The Quarterly Journal of Economics* **110**(4), 1127.

Islam, N. (2003), 'What have we learnt from the convergence debate?', *Journal of economic surveys* **17**(3), 309–362.

Klenow, P.J. and A. Rodriguez-Clare (2005), 'Externalities and growth', *Handbook of Economic Growth* **1**, 817–861.

Koop, G., J. Osiewalski and M.F.J. Steel (2000), 'Modeling the sources of output growth in a panel of countries', *Journal of Business & Economic Statistics* pp. 284–299.

Koopmans, T.C. (1965), 'On the concept of optimal growth', *The econometric approach to development planning* pp. 225–287.

Lee, K., M.H. Pesaran and R.P. Smith (1997), 'Growth and convergence in a multi-country empirical stochastic solow model', *Journal of Applied Econometrics* **12**(4), 357–392.

Lee, L. and J. Yu (2010), 'A spatial dynamic panel data model with both time and individual fixed effects', *Econometric Theory* **26**(02), 564–597.

Lee, L.F. (1983), 'Generalized econometric models with selectivity', *Econometrica: Journal of the Econometric Society* pp. 507–512.

Lee, L.F. (2002), 'Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models', *Econometric theory* **18**(02), 252–277.

Lee, L.F. (2004), 'Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models', *Econometrica* pp. 1899–1925.

Lehmann, E.L. and J.P. Romano (2005), *Testing statistical hypotheses*, Springer Verlag.

Mankiw, N.G., D. Romer and D.N. Weil (1992), 'A contribution to the empirics of economic growth', *The Quarterly Journal of Economics* **107**(2), 407.

Nerlove, M. (1971), 'Further evidence on the estimation of dynamic economic relations from a time series of cross sections', *Econometrica: Journal of the Econometric Society* pp. 359–382.

Nerlove, M. (1999), *Properties of alternative estimators of dynamic panel models: an empirical analysis of cross-country data for the study of economic growth*, Cambridge: Cambridge University Press.

Phillips, P.C.B., D. Sul and Yale University. Cowles Foundation for Research in Economics (2003), *The elusive empirical shadow of growth convergence*, Cowles Foundation for Research in Economics.

Quah, D., London School of Economics and Political Science. Financial Markets Group (1992), *Empirical cross-section dynamics in economic growth*, Citeseer.

Rodriguez, F. and D. Rodrik (2001), 'Trade policy and economic growth: a skeptic's guide to the cross-national evidence'.

Saikkonen, P. (1989), 'Asymptotic relative efficiency of the classical test statistics under misspecification', *Journal of Econometrics* **42**(3), 351–369.

Sala-i Martin, X.X. (1996), 'The classical approach to convergence analysis', *The Economic Journal* pp. 1019–1036.

Sen, M. and A.K. Bera (2011), 'Specification testing for panel spatial models', Working Paper.

Serfling, R.J. (1980), ‘Approximation theorems of mathematical statistics’, *New York* .

Solow, R.M. (1956), ‘A contribution to the theory of economic growth’, *The Quarterly Journal of Economics* **70**(1), 65.

Temple, J.R.W. (1998), ‘Robustness tests of the augmented solow model’, *Journal of Applied Econometrics* **13**(4), 361–375.

Yu, J. and L. Lee (2009), ‘Convergence: A spatial dynamic panel data approach’, *Submitted to Global Journal of Economics, in July* **23**.

8 Appendix

We consider the following dynamic panel spatial model which is the combination of all the different piecewise framework which has been discussed in Section 2.

$$y_{it} = \gamma y_{it-1} + \tau \sum_{j=1}^{N} m_{ij} y_{jt} + \delta \sum_{j=1}^{N} m_{ij} y_{jt-1} + X_{it}\beta + u_{it} \tag{51}$$

$$u_{it} = \mu_i + \epsilon_{it} \tag{52}$$

$$\epsilon_{it} = \lambda \sum_{j=1}^{N} w_{ij} \epsilon jt + v_{it} \tag{53}$$

$$v_{it} = \rho v_{it-1} + e_{it}, \text{where } e_{it} \sim IIDN(0, \sigma_e^2) \tag{54}$$

for $i = 1, 2, \ldots, N; t = 1, 2, \ldots, T$. Here y_{it} is the observation for i^{th} individual/observation at t^{th} time, X_{it} denotes the observations on non-stochastic regressors and u_{it} is the regression disturbance. Spatial dependence is captured by the weigh matrices $M = (m_{ij})$ and $W = (w_{ij})$. In this framework, I have considered spatial lag dependence (τ), time dynamics (γ), space recursive (δ), spatial error dependence (λ), serial correlation in error (ρ) and individual effect (μ_i). I consider random effect model here, i.e., $\mu_i \sim IID(0, \sigma_\mu)$, like Sen and Bera (2011). W and M are row-standardized weight matrices whose diagonal elements are zero, such that $(I - \tau M)$ and $(I - \lambda W)$ are non-singular, where I is an identity matrix of dimension N. I assume that the model satisfies the regularity conditions given in Lee and Yu (2010).

In matrix form, the equations (9) - (12) can be rewritten as

$$y = \tau(I_T \otimes M)y + [(\gamma + \delta M) \otimes I_T]ly + X\beta + u \tag{55}$$

where y is f dimension $NT \times 1, X$ is $NT \times K, \beta$ is $k \times 1$ and u is $NT \times 1$. Here l is the lag operator, X is assumed to be of full column rank and its elements are bounded in absolute value. The disturbance term can be expressed as

$$u = (\iota_T \otimes I_N)\mu + (I_T \otimes B^{-1})v \tag{56}$$

where $B = (I_N - \lambda W), \iota_T$ is vector of ones of dimension T, I_T is an identity matrix of dimension $T \times T$ and $\otimes$ denotes Kronecker product. Under this setup, the variance-covariance matrix of u is given by

$$\Omega = \sigma_{\mu^2}[J_T \otimes I_N] + [V \otimes (B'B)^{-1}] \tag{57}$$

where J_T is a matrix of ones of dimension $T \times T$, and V is the familiar $T \times T$ variance -covariance matrix for AR (1) process in equation (20), i.e.,

$$V = E(v'v) = [\frac{1}{1-\rho^2}V_1] \otimes \sigma_e^2 I_N = V_\rho \otimes \sigma_e^2 I_N \tag{58}$$

with

$$V_1 = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{T-1} \\ \vdots & \vdots & \ddots & \vdots & \\ \rho^{T-1} & \rho^{T-2} & \dots & 1 & \end{bmatrix}$$

and $V_\rho = \frac{1}{1-\rho^2}V_1$.

The loglikelihood function of the above model can be written as:

$$L = \frac{-NT}{2}ln2\pi - \frac{1}{2}\ln|\Omega| + T\ln|A| - \frac{1}{2}[(I_T \otimes A)y - [(\gamma+\delta M)\otimes I_T]ly - X\beta]'\Omega^{-1}[(I_T \otimes A)y - [(\gamma+\delta M)\otimes I_T]ly - X\beta] \tag{59}$$

where $A = (I_N - \tau M)$ and following Sen and Bera (2011), I can write

$$\frac{1}{2}\ln|\Omega| = -\frac{N}{2}\ln(1-\rho^2) + \frac{1}{2}\ln|d^2(1-\rho)^2\phi I_N + (B'B)^{-1}| + \frac{NT}{2}\ln\sigma_e^2 - (T-1)\ln|B|$$

where $d^2 = \alpha^2 + (T-1)$, $\alpha = \sqrt{\frac{1+\rho}{1-\rho}}$ and $\phi = \frac{\sigma_\mu^2}{\sigma_e^2}$. Thus substituting $\frac{1}{2}\ln|\Omega|$ in L , I obtain

$$\begin{aligned} L = &\frac{-NT}{2}ln2\pi + \frac{N}{2}\ln(1-\rho^2) - \frac{1}{2}\ln|d^2(1-\rho)^2\phi I_N + (B'B)^{-1}| - \frac{NT}{2}\ln\sigma_e^2 + (T-1)\ln|B| + T\ln|A| \\ &- \frac{1}{2}[(I_T \otimes A)y - [(\gamma+\delta M)\otimes I_T]ly - X\beta]'\Omega^{-1}[(I_T \otimes A)y - [(\gamma+\delta M)\otimes I_T]ly - X\beta] \end{aligned} \tag{60}$$

8.1 Derivation of Score

$$\frac{\partial L}{\partial \beta} = X'\Omega^{-1}u \tag{61}$$

$$\frac{\partial L}{\partial \sigma_e^2} = -\frac{1}{2}trC^{-1}\frac{(d^2(1-\rho)^2\sigma_\mu^2 I_N}{\sigma_e^4} - \frac{NT}{2\sigma_e^2} - \frac{1}{2}u'(\Omega^{-1}(V_\rho \otimes [(B'B)^{-1}])\Omega^{-1})u \tag{62}$$

$$\frac{\partial L}{\partial \gamma} = (I_T \otimes Y_{NT-1})'\Omega^{-1}u \tag{63}$$

$$\frac{\partial L}{\partial \sigma_\mu^2} = -\frac{1}{2}trC^{-1}\frac{(d^2(1-\rho)^2 I_N}{\sigma_e^2} + \frac{1}{2}u'\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}u \tag{64}$$

$$\frac{\partial L}{\partial \rho} = -\frac{N\rho}{1-\rho^2} + \frac{1}{2}trC^{-1}(\rho + (T-1)(1-\rho)\phi I_N) + \frac{1}{2}u'(\sigma_e^{-2}(\frac{1}{1-\rho^2})^2[2\rho V_1 + (1-\rho^2)F_\rho] \otimes (B'B)^{-1})u \tag{65}$$

$$\frac{\partial L}{\partial \delta} = [(W \otimes I_T)Y_{NT-1}]\Omega^{-1}u \tag{66}$$

$$\frac{\partial L}{\partial \tau} = -Ttr(A^{-1}W) + \frac{1}{2}\Omega^{-1}(I_T \otimes W)y \tag{67}$$

$$\frac{\partial L}{\partial \lambda} = -(T-1)tr(B^{-1}W) + \frac{1}{2}trC^{-1}[(B'B)^{-1}[B'W + W'B](B'B)^{-1}] - \frac{1}{2}u'\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}u \tag{68}$$

where $C = (d^2(1-\rho)^2\phi I_N + (B'B)^{-1})$. The score functions evaluated under H_0^a, i.e., restricted MLE of θ_0 with $\tilde{\omega} = (\tilde{\beta}, \tilde{\sigma_e^2})$ are:

$$\frac{\partial L}{\partial \beta} = 0 \tag{69}$$

$$\frac{\partial L}{\partial \sigma_e^2} = 0 \tag{70}$$

$$\frac{\partial L}{\partial \gamma} = \frac{[I_T \otimes Y_{NT-1}]\tilde{u}'}{\sigma_e^2} \tag{71}$$

$$\frac{\partial L}{\partial \sigma_\mu^2} = \frac{NT}{2\tilde{\sigma_e^2}}[\frac{\tilde{u}'(J_T \otimes I_N)\tilde{u}}{\tilde{u}'\tilde{u}} - 1] \tag{72}$$

$$\frac{\partial L}{\partial \rho} = \frac{NT}{2}[\frac{\tilde{u}'(G \otimes I_N)\tilde{u}}{\tilde{u}'\tilde{u}}] \tag{73}$$

$$\frac{\partial L}{\partial \delta} = \frac{\tilde{u}'[(I_T \otimes W)Y_{NT-1}]}{\tilde{\sigma_e^2}} \tag{74}$$

$$\frac{\partial L}{\partial \tau} = \frac{\tilde{u}'[(I_T \otimes W)Y_{NT}]}{\tilde{\sigma_e^2}} \tag{75}$$

$$\frac{\partial L}{\partial \lambda} = \frac{NT}{2}[\frac{\tilde{u}'(I_T \otimes (W + W'))\tilde{u}}{\tilde{u}'\tilde{u}}] \tag{76}$$

where $\tilde{u} = y - x\tilde{\beta}$ is the OLS residual vector, and $\tilde{\sigma_e^2} = \frac{\tilde{u}'\tilde{u}}{NT}$.

8.2 Derivation of Information Matrix

To derive the information matrix under joint null H_o^a, I need to derive the second order derivatives and take expectation. Differentiating equation (69) wrt $\beta, \sigma_e^2, \gamma, \sigma_\mu^2, \rho, \delta, \tau$ and λ, we get

$$\frac{\partial^2 L}{\partial \beta \partial \beta'} = -X'\Omega^{-1}X \tag{77}$$

$$\frac{\partial^2 L}{\partial \beta \partial \sigma_e^2} = -u'\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}X \tag{78}$$

$$\frac{\partial^2 L}{\partial\beta\partial\gamma} = -(I_T \otimes Y_{NT-1})'\Omega^{-1}X \tag{79}$$

$$\frac{\partial^2 L}{\partial\beta\partial\sigma_\mu^2} = -X'\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}u \tag{80}$$

$$\frac{\partial^2 L}{\partial\beta\partial\rho} = -X'\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}u \tag{81}$$

$$\frac{\partial^2 L}{\partial\beta\partial\delta} = -[(W \otimes I_T)Y_{NT-1}]'\Omega^{-1}X \tag{82}$$

$$\frac{\partial^2 L}{\partial\beta\partial\tau} = -X'\Omega^{-1}(I_T \otimes W)Y \tag{83}$$

$$\frac{\partial^2 L}{\partial\beta\partial\lambda} = -X'\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}u \tag{84}$$

Differentiating equation (70) wrt $\gamma, \sigma_\mu^2, \rho, \delta, \tau$ and λ, we get

$$\frac{\partial^2 L}{\partial\sigma_e^2\partial\gamma} = -(I_T \otimes Y_{NT-1})\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}u \tag{85}$$

$$\frac{\partial^2 L}{\partial\sigma_e^2\partial\sigma_\mu^2} = \frac{1}{2}tr[C^{-1}\frac{d^2(1-\rho)^2 I_N}{\sigma_e^2}C^{-1}(B'B)^{-1}] - u\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}u \tag{86}$$

$$\begin{aligned}\frac{\partial^2 L}{\partial\sigma_e^2\partial\rho} &= -tr[C^{-1}[(\rho + (1-\rho)(T-1))\phi I_N]C^{-1}(B'B)^{-1}] \\ &\quad - u\Omega^{-1}[[2\frac{\rho}{1-\rho^2}^2 V_1 + \frac{1}{1-\rho^2}F_\rho] \otimes (B'B)^{-1}]\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}u \\ &\quad + \frac{1}{2}u\Omega^{-1}[[\frac{2\rho}{(1-\rho^2)^2}V_1 + \frac{1}{1-\rho^2}F_\rho] \otimes (B'B)^{-1}]\end{aligned} \tag{87}$$

$$\frac{\partial^2 L}{\partial\sigma_e^2\partial\delta} = -((W \otimes I_T)Y_{NT-1})'\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}u \tag{88}$$

$$\frac{\partial^2 L}{\partial\sigma_e^2\partial\tau} = -((W \otimes I_T)Y_{NT})'\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}u \tag{89}$$

$$\begin{aligned}\frac{\partial^2 L}{\partial\sigma_e^2\partial\lambda} &= -\frac{1}{2}tr[C^{-1}(B'B)^{-1}(B'W+W'B)(B'B)^{-1}C^{-1}(B'B)^{-1}+C^{-1}(B'B)^{-1}(B'W+W'B)(B'B)^{-1}] \\ &-\frac{T-1}{2}tr[(B'B)^{-1}(B'W+W'B)(B'B)^{-1}]-u\Omega^{-1}[V_\rho\otimes(B'B)^{-1}[B'W+W'B](B'B)^{-1}]\Omega^{-1}(V_\rho\otimes(B'B)^{-1})\Omega^{-1}u \\ &\quad + \frac{1}{2}u\Omega^{-1}(V_\rho \otimes (B'B)^{-1}(B'W+W'B)(B'B)^{-1})\Omega^{-1}u\end{aligned} \tag{90}$$

Differentiating equation (71) wrt $\gamma, \sigma_\mu^2, \rho, \delta, \tau$ and λ, we get

$$\frac{\partial^2 L}{\partial\gamma\partial\gamma} = -(I_T \otimes Y_{NT-1})'\Omega^{-1}(I_T \otimes Y_{NT-1}) \tag{91}$$

$$\frac{\partial^2 L}{\partial\gamma\partial\sigma_\mu^2} = -(I_T \otimes Y_{NT-1})'\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}u \tag{92}$$

$$\frac{\partial^2 L}{\partial\gamma\partial\rho} = -(I_T \otimes Y_{NT-1})'\Omega^{-1}[[\frac{2\rho}{(1-\rho^2)^2}V_1 + \frac{1}{(1-\rho^2)}F_\rho] \otimes (B'B)^{-1}]\Omega^{-1}u \tag{93}$$

$$\frac{\partial^2 L}{\partial\gamma\partial\delta} = -(I_T \otimes Y_{NT-1})'\Omega^{-1}[(W \otimes I_T)Y_{NT-1}] \tag{94}$$

$$\frac{\partial^2 L}{\partial\gamma\partial\tau} = -(I_T \otimes Y_{NT-1})'\Omega^{-1}[(W \otimes I_T)Y_{NT}] \tag{95}$$

$$\frac{\partial^2 L}{\partial\gamma\partial\lambda} = -(I_T \otimes Y_{NT-1})'\Omega^{-1}[V_\rho \otimes (B'B)^{-1}(B'W + W'B)(B'B)^{-1}]\Omega^{-1}u \tag{96}$$

Differentiating equation (72) wrt $\sigma_\mu^2, \rho, \delta, \tau$ and λ, we get

$$\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\sigma_\mu^2} = \frac{1}{2}tr[C^{-1}\frac{d^2(1-\rho)^2 I_N}{\sigma_e^2}C^{-1}\frac{d^2(1-\rho)^2 I_N}{\sigma_e^2}] - u\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}u \tag{97}$$

$$\begin{aligned}\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\rho} = \frac{1}{2}tr[\frac{d^2(1-\rho)^2 I_N}{\sigma_e^2}[C^{-1}(\rho + (T-1)(1-\rho)\phi I_N C^{-1})] + u'\Omega^{-1}[\sigma_e^2(\frac{1}{1-\rho^2})^2[2\rho V_1 \\ + (1-\rho^2)F_\rho] \otimes (B'B)^{-1}]\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}u\end{aligned} \tag{98}$$

$$\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\delta} = -[(W \otimes I_T)Y_{NT-1}]'\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}u \tag{99}$$

$$\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\tau} = -[(W \otimes I_T)Y_{NT}]'\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}u \tag{100}$$

$$\begin{aligned}\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\lambda} = -\frac{1}{2}tr[\frac{d^2(1-\rho)^2}{\sigma_e^2}I_N C^{-1}\frac{d^2(1-\rho)^2}{\sigma_e^2}I_N C^{-1}] \\ + u'\Omega^{-1}[\sigma_e^2(V_\rho \otimes [(B'B)^{-1}[B'W + W'B](B'B)^{-1}]\Omega^{-1}(J_T \otimes I_N)\Omega^{-1}u\end{aligned} \tag{101}$$

Differentiating equation (73) wrt ρ, δ, τ and λ, we get

$$\frac{\partial^2 L}{\partial\rho\partial\rho} = \frac{-N + N\rho^2}{(1-\rho^2)^2} + \frac{1}{2}tr((2\rho + (T-1)(1-\rho))\phi I_N)[C^{-1}(\rho + (T-1)(1-\rho))\phi I_N C^{-1}]$$
$$+ u\Omega^{-1}[\sigma_e^2(\frac{1}{1-\rho^2})^2[2\rho V_1 + (1-\rho^2)F_\rho] \otimes (B'B)^{-1}]\Omega^{-1}[\sigma_e^2(\frac{1}{1-\rho^2})^2[2\rho V_1$$
$$+(1-\rho^2)F_\rho]\otimes(B'B)^{-1}]\Omega^{-1}u + u\Omega^{-1}[[\sigma_e^2(\frac{1}{1-\rho^2})^2[2V_1 - 2\rho F_\rho] + 4(1-\rho^2)\rho(2\rho V_1 + (1-\rho^2)F_\rho)\otimes(B'B)^{-1}]\Omega^{-1}u \tag{102}$$

$$\frac{\partial^2 L}{\partial\rho\partial\delta} = -[(W \otimes I_T)Y_{NT-1}]\Omega^{-1}[[\frac{2\rho}{(1-\rho^2)^2}V_1 + \frac{1}{1-\rho^2}F_\rho] \otimes (B'B)^{-1}]\Omega^{-1}u \tag{103}$$

$$\frac{\partial^2 L}{\partial\rho\partial\tau} = -[(W \otimes I_T)Y_{NT}]\Omega^{-1}[[\frac{2\rho}{(1-\rho^2)^2}V_1 + \frac{1}{1-\rho^2}F_\rho] \otimes (B'B)^{-1}]\Omega^{-1}u \tag{104}$$

$$\frac{\partial^2 L}{\partial\rho\partial\lambda} = \frac{1}{2}tr[[(B'B)^{-1}[B'W + W'B](B'B)^{-1}C6{-}1(\rho + (T-1)(1-\rho))\phi I_N C^{-1}]$$
$$+ u\Omega^{-1}[\sigma_e^2(\frac{1}{1-\rho^2})^2[2V_1 - 2\rho F_\rho] \otimes (B'B)^{-1}]\Omega^{-1}[\sigma_e^2(V_\rho \otimes [(B'B)^{-1}[B'W + W'B](B'B)^{-1}]\Omega^{-1} -$$
$$\frac{1}{2}u\Omega^{-1}[\sigma_e^4(\frac{1}{1-\rho^2})^2[2V_1 - 2\rho F_\rho] \otimes [(B'B)^{-1}[B'W + W'B](B'B)^{-1}](B'B)^{-1}]\Omega^{-1}u \tag{105}$$

Differentiating equation (74) wrt δ, τ and λ, we get

$$\frac{\partial^2 L}{\partial\delta\partial\delta} = -[(W \otimes I_T)Y_{NT-1}]'\Omega^{-1}[(W \otimes I_T)Y_{NT-1}] \tag{106}$$

$$\frac{\partial^2 L}{\partial\delta\partial\tau} = -[(W \otimes I_T)Y_{NT-1}]'\Omega^{-1}[(W \otimes I_T)Y_{NT}] \tag{107}$$

$$\frac{\partial^2 L}{\partial\delta\partial\lambda} = -[(W \otimes I_T)Y_{NT-1}]'\Omega^{-1}(V_\rho(B'B)^{-1}(W'B + B'W)(B'B)^{-1})\Omega^{-1}u \tag{108}$$

Differentiating equation (75) wrt τ and λ, we get

$$\frac{\partial^2 L}{\partial\tau\partial\tau} = -Ttr((A^{-1}W)^2) - Y_{NT}(I_T \otimes W)\Omega^{-1}(I_T \otimes W)Y_{NT} \tag{109}$$

$$\frac{\partial^2 L}{\partial\tau\partial\lambda} = -u\Omega^{-1}(V_\rho \otimes (B'B)^{-1})\Omega^{-1}(I_T \otimes W)u \tag{110}$$

For $\frac{\partial^2 L}{\partial\lambda\partial\lambda}$ please see Sen and Bera 2011, Technical Appendix.

Under the joint null $H_o^a : \gamma = \sigma_\mu^2 = \rho = \delta = \tau = \lambda = 0$, the non-zero second-order derivatives are :

$\frac{\partial^2 L}{\partial\beta\partial\beta} = -\frac{X'X}{\hat{\sigma_e^2}}$

$$\frac{\partial^2 L}{\partial\beta\partial\gamma} = -\frac{(I_T \otimes Y_{NT-1})' X}{\hat{\sigma_e^2}}$$

$$\frac{\partial^2 L}{\partial\beta\partial\delta} = -\frac{[(W \otimes I_T) Y_{NT-1}]' X}{\hat{\sigma_e^2}}$$

$$\frac{\partial^2 L}{\partial\beta\partial\tau} = -\frac{X'(W \otimes I_T) X\hat{\beta}}{\hat{\sigma_e^2}}$$

$$\frac{\partial^2 L}{\partial\sigma_e^2\partial\sigma_e^2} = -\frac{NT}{2\hat{\sigma_e^4}}$$

$$\frac{\partial^2 L}{\partial\sigma_e^2\partial\gamma} = -\frac{(I_T \otimes Y_{NT-1})(I_T \otimes I_N) u}{2\hat{\sigma_e^4}}$$

$$\frac{\partial^2 L}{\partial\sigma_e^2\partial\sigma_\mu^2} = -\frac{NT}{2\hat{\sigma_e^4}}$$

$$\frac{\partial^2 L}{\partial\gamma\partial\gamma} = -\frac{(I_T \otimes Y_{NT-1})'(I_T \otimes Y_{NT-1})}{\hat{\sigma_e^2}}$$

$$\frac{\partial^2 L}{\partial\gamma\partial\sigma_\mu^2} = -\frac{(I_T \otimes Y_{NT-1})'(J_T \otimes I_N) u}{2\hat{\sigma_e^4}}$$

$$\frac{\partial^2 L}{\partial\gamma\partial\rho} = -\frac{(I_T \otimes Y_{NT-1})'(I_T \otimes I_N) u}{2\hat{\sigma_e^4}}$$

$$\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\sigma_\mu^2} = -\frac{NT^2}{2\hat{\sigma_e^4}}$$

$$\frac{\partial^2 L}{\partial\sigma_\mu^2\partial\rho} = -\frac{N(T-1)}{\hat{\sigma_e^2}}$$

$$\frac{\partial^2 L}{\partial\rho\partial\rho} = -N(T-1)$$

$$\frac{\partial^2 L}{\partial\delta\partial\delta} = -\frac{((W \otimes I_T) Y_{NT-1})'((W \otimes I_T) Y_{NT-1})}{\hat{\sigma_e^2}}$$

$$\frac{\partial^2 L}{\partial\delta\partial\tau} = -\frac{((W \otimes I_T) Y_{NT-1})'((W \otimes I_T) Y_{NT-1})}{\hat{\sigma_e^2}}$$

$$\frac{\partial^2 L}{\partial\delta\partial\lambda} = -\frac{((W \otimes I_T) Y_{NT-1})'((I_T \otimes (W+W'))) u}{\hat{\sigma_e^2}}$$

$$\frac{\partial^2 L}{\partial\tau\partial\tau} = -(Ttr(W^2 + WW') + \frac{(\hat{\beta}' X'(I_T \otimes W')(I_T \otimes W) X\hat{\beta})}{\hat{\sigma_e^2}})$$

$$\frac{\partial^2 L}{\partial\tau\partial\lambda} = \frac{\partial^2 L}{\partial\lambda\partial\lambda} = -Ttr(W^2 + WW').$$

All the other second derivatives becomes zero under joint null. Thus the information matrix J, equation (12), under H_o^a is

$$J(\theta_o) = \begin{bmatrix} J_\beta & 0 & J_{\beta\gamma} & 0 & 0 & J_{\beta\delta} & J_{\beta\tau} & 0 \\ 0 & J_{\sigma_e^2} & J_{\sigma_e^2\gamma} & J_{\sigma_e^2\sigma_\mu^2} & 0 & 0 & 0 & 0 \\ J_{\gamma\beta} & J_{\gamma\sigma_e^2} & J_\gamma & J_{\gamma\sigma_\mu^2} & J_{\gamma\rho} & 0 & 0 & 0 \\ 0 & J_{\sigma_\mu^2\sigma_e^2} & J_{\sigma_\mu^2\gamma} & J_{\sigma_\mu^2} & J_{\sigma_\mu^2\rho} & 0 & 0 & 0 \\ 0 & 0 & J_{\rho\gamma} & J_{\rho\sigma_\mu^2} & J_\rho & 0 & 0 & 0 \\ J_{\delta\beta} & 0 & 0 & 0 & 0 & J_\delta & J_{\delta\tau} & J_{\delta\lambda} \\ J_{\tau\beta} & 0 & 0 & 0 & 0 & J_{\tau\delta} & J_\tau & J_{\tau\lambda} \\ 0 & 0 & 0 & 0 & 0 & J_{\lambda\delta} & J_{\lambda\tau} & J_\lambda \end{bmatrix} \quad (111)$$

where $J = E(-\frac{1}{NT}\frac{\partial^2 L}{\partial\tilde{\theta}\partial\tilde{\theta}'})$ evaluated at θ_0.

8.3 Derivation of test statistics

Recall from Section 3:

$$RS^*_\psi = \frac{1}{N}[d_\psi(\tilde{\theta}) - J_{\psi\phi.\omega}(\tilde{\theta})J^{-1}_{\phi.\omega}(\tilde{\theta})d_\phi(\tilde{\theta})][J_{\psi.\omega}(\tilde{\theta}) - J_{\psi\phi.\omega}(\tilde{\theta})J^{-1}_{\phi.\omega}(\tilde{\theta})J_{\phi\psi.\omega}(\tilde{\theta})]^{-1}[d_\psi(\tilde{\theta}) - J_{\psi\phi.\omega}(\tilde{\theta})J^{-1}_{\phi.\omega}(\tilde{\theta})d_\phi(\tilde{\theta})]' \tag{112}$$

where $\omega = (\beta', \sigma_e^2)', \psi$ and ϕ are different combinations of the parameters $(\gamma, \sigma^2_\mu, \rho, \delta, \tau, \lambda)$.

I) $H^b_o : \gamma = 0$ in presence of $\phi = (\sigma^2_\mu, \rho, \delta, \tau, \lambda)$.

Here we are testing the significance of time-dynamics γ, in presence of random effects, serial correlation, and spatial dependence.

$d_\psi = d_\gamma$

$d_\phi = (d_{\sigma^2_\mu}, d_\rho, d_\delta, d_\tau, d_\lambda)$

$J_{\psi\phi.\omega} = J_{\psi\phi} - J_{\psi\omega}J^{-1}_\omega J_{\phi\omega} = (J_{\gamma\sigma^2_\mu.\sigma^2_e}, J_{\gamma\rho}, 0, 0, 0)$

$J_{\phi.\omega} = J_\phi - J_{\phi\omega}J^{-1}_\omega J_{\omega\phi} =$

$$\begin{bmatrix} J_{\sigma^2_\mu.\sigma^2_e} & J_{\sigma^2_\mu\rho} & 0 & 0 & 0 \\ J_{\rho\sigma^2_\mu} & J_\rho & 0 & 0 & 0 \\ 0 & 0 & J_{\delta.\beta} & J_{\delta\tau.\beta} & J_{\delta\lambda} \\ 0 & 0 & J_{\tau\delta.\beta} & J_{\tau.\beta} & J_{\tau\lambda} \\ 0 & 0 & J_{\lambda\delta} & J_{\lambda\tau} & J_\lambda \end{bmatrix}$$

Therefore, adjusted proposed test statistic for time-dynamics γ is:

$$RS^*_\gamma = [d_\gamma - J_{\gamma\sigma^2_\mu.\sigma^2_e}J^{-1}_{\sigma^2_\mu.\sigma^2_e}d_{\sigma^2_\mu} - J_{\gamma\rho}J^{-1}_\rho d_\rho][J_{\gamma.\omega} - J_{\gamma\sigma^2_\mu.\sigma^2_e}J^{-1}_{\sigma^2_\mu.\sigma^2_e}J_{\sigma^2_\mu\gamma.\sigma^2_e} - J_{\gamma\rho}J^{-1}_\rho J_{\rho\gamma}]^{-1}$$
$$[d_\gamma - J_{\gamma\sigma^2_\mu.\sigma^2_e}J^{-1}_{\sigma^2_\mu.\sigma^2_e}d_{\sigma^2_\mu} - J_{\gamma\rho}J^{-1}_\rho d_\rho]' \to \chi^2_1(0). \tag{113}$$

II) $H^c_o : \sigma^2_\mu = 0$ in presence of $\gamma, \rho, \delta, \tau, \lambda$.

Here $\phi = (\gamma, \rho, \delta, \tau, \lambda)$

$d_\psi = d_{\sigma_{\mu^2}}$

$d_\phi = (d_\gamma, d_\rho, d_\delta, d_\tau, d_\lambda)$

$J_{\psi\phi.\omega} = (J_{\sigma^2_\mu\gamma.\sigma^2_e}, J_{\sigma^2_\mu\rho}, 0, 0, 0)$

$$J_{\phi.\omega} = \begin{bmatrix} J_{\gamma.\omega} & J_{\gamma\rho} & 0 & 0 & 0 \\ J_{\rho\gamma} & J_\rho & 0 & 0 & 0 \\ 0 & 0 & J_{\delta.\beta} & J_{\delta\tau.\beta} & J_{\delta\lambda} \\ 0 & 0 & J_{\tau\delta.\beta} & J_{\tau.\beta} & J_{\tau\lambda} \\ 0 & 0 & J_{\lambda\delta} & J_{\lambda\tau} & J_\lambda \end{bmatrix}$$

The adjusted RS test statistics is:

$$RS^*_{\sigma^2_\mu} = [d_{\sigma^2_\mu} - J_{\sigma^2_\mu\gamma.\sigma^2_e}J^{-1}_{\gamma.\omega}d_\gamma - J_{\sigma^2_\mu\rho}J^{-1}_\rho d_\rho][J_{\sigma^2_\mu.\sigma^2_e} - J_{\sigma^2_\mu\gamma.\sigma^2_e}J^{-1}_{\gamma.\omega}J_{\gamma\sigma^2_\mu.\sigma^2_e} - J_{\sigma^2_\mu\rho}J^{-1}_\rho J_{\rho\sigma^2_\mu}]^{-1}$$
$$[d_{\sigma^2_\mu} - J_{\sigma^2_\mu\gamma.\sigma^2_e}J^{-1}_{\gamma.\omega}d_\gamma - J_{\sigma^2_\mu\rho}J^{-1}_\rho d_\rho]' \to \chi^2_1(0), \quad (114)$$

III) $H^d_o : \rho = 0$ in presence of $\gamma, \sigma^2_\mu, \delta, \tau, \lambda$.

Here $\phi = (\gamma, \sigma^2_\mu, \delta, \tau, \lambda)$.
$d_\psi = d_\rho$
$d_\phi = (d_\gamma, d_{\sigma_{\mu^2}}, d_\delta, d_\tau, d_\lambda)$
$J_{\psi\phi.\omega} = (J_{\rho\gamma}, J_{\rho\sigma^2_\mu}, 0, 0, 0)$.

$$J_{\phi.\omega} = \begin{bmatrix} J_{\gamma.\omega} & J_{\gamma\sigma^2_\mu.\sigma^2_e} & 0 & 0 & 0 \\ J_{\sigma^2_\mu\gamma.\sigma^2_e} & J_{\sigma^2_\mu.\sigma^2_e} & 0 & 0 & 0 \\ 0 & 0 & J_{\delta.\beta} & J_{\delta\tau.\beta} & J_{\delta\lambda} \\ 0 & 0 & J_{\tau\delta.\beta} & J_{\tau.\beta} & J_{\tau\lambda} \\ 0 & 0 & J_{\lambda\delta} & J_{\lambda\tau} & J_\lambda \end{bmatrix}$$

The adjusted test statistic is:

$$RS^*_\rho = [d_\rho - J_{\rho\gamma.\sigma^2_e}J^{-1}_{\gamma.\omega}d_\gamma - J_{\rho\sigma^2_\mu.\sigma^2_e}J^{-1}_{\sigma^2_\mu.\sigma^2_e}d_{\sigma^2_\mu}][J_\rho - J_{\rho\gamma.\sigma^2_e}J^{-1}_{\gamma.\omega}J_{\gamma\rho.\sigma^2_e} - J_{\rho\sigma^2_\mu.\sigma^2_e}J^{-1}_{\sigma^2_\mu.\sigma^2_e}J_{\sigma^2_\mu\rho.\sigma^2_e}]^{-1}$$
$$[d_\rho - J_{\rho\gamma.\sigma^2_e}J^{-1}_{\gamma.\omega}d_\gamma - J_{\rho\sigma^2_\mu.\sigma^2_e}J^{-1}_{\sigma^2_\mu.\sigma^2_e}d_{\sigma^2_\mu}]' \to \chi^2_1 \quad (115)$$

IV) $H^e_o : \delta = 0$ in presence of $\gamma, \sigma^2_\mu, \rho, \tau, \lambda$.

Here $\phi = (\gamma, \sigma^2_\mu, \rho, \tau, \lambda)$.
$d_\psi = d_\delta$
$d_\phi = (d_\gamma, d_{\sigma_{\mu^2}}, d_\rho, d_\tau, d_\lambda)$
$J_{\psi\phi.\omega} = (0, 0, 0, J_{\delta\lambda}, J_{\delta\tau.\beta})$.

$$J_{\phi.\omega} = \begin{bmatrix} J_{\gamma.\omega} & J_{\gamma\sigma^2_\mu.\sigma^2_e} & J_{\gamma\rho} & 0 & 0 \\ J_{\sigma^2_\mu\gamma.\sigma^2_e} & J_{\sigma^2_\mu.\sigma^2_e} & J_{\sigma^2_\mu\rho} & 0 & 0 \\ J_{\rho\gamma} & J_{\rho\sigma^2_\mu} & J_\rho & 0 & 0 \\ 0 & 0 & 0 & J_{\tau.\beta} & J_{\tau\lambda} \\ 0 & 0 & 0 & J_{\lambda\tau} & J_\lambda \end{bmatrix}$$

The test statistic for space recursive parameter δ is directly affected by the other spatial parameters λ and τ, and not by other parameters. The separation between the spatial parameters and all the other parameters is also distinct here, as far as the test statistic is concerned. The adjusted RS test statistic is:

$$RS^*_\delta = [d_\delta - J_{(\delta\lambda).(\tau.\beta)} J^{-1}_{\lambda.(\tau.\beta)} d_\lambda - J_{(\delta\tau.\beta).\lambda} J^{-1}_{(\tau.\beta).\lambda} d_\tau][J_\delta - J_{(\delta\lambda).(\tau.\beta)} J^{-1}_{\lambda.(\tau.\beta)} J_{(\lambda\delta).(\tau.\beta)} - J_{(\delta\tau.\beta).\lambda} J^{-1}_{(\tau.\beta).\lambda} J_{(\tau\delta.\beta).\lambda}]^{-1}$$
$$[d_\delta - J_{(\delta\lambda).(\tau.\beta)} J^{-1}_{\lambda.(\tau.\beta)} d_\lambda - J_{(\delta\tau.\beta).\lambda} J^{-1}_{(\tau.\beta).\lambda} d_\tau]' \sim \chi^2_1(0) \quad (116)$$

V) $H_o^f : \tau = 0$ in presence of $\gamma, \sigma^2_\mu, \rho, \delta, \lambda$.

Here $\phi = (\gamma, \sigma^2_\mu, \rho, \delta, \lambda)$
$d_\psi = d_\tau$
$d_\phi = (d_\gamma, d_{\sigma_{\mu^2}}, d_\rho, d_\delta, d_\lambda)$
$J_{\psi\phi.\omega} = (0, 0, 0, J_{\tau\delta.\beta}, J_{\tau\lambda})$.

$$J_{\phi.\omega} \begin{bmatrix} J_{\gamma.\omega} & J_{\gamma\sigma^2_\mu.\sigma^2_e} & J_{\gamma\rho} & 0 & 0 \\ J_{\sigma^2_\mu\gamma.\sigma^2_e} & J_{\sigma^2_\mu.\sigma^2_e} & J_{\sigma^2_\mu\rho} & 0 & 0 \\ J_{\rho\gamma} & J_{\rho\sigma^2_\mu} & J_\rho & 0 & 0 \\ 0 & 0 & 0 & J_{\delta.\tau} & J_{\delta\lambda} \\ 0 & 0 & 0 & J_{\lambda\delta} & J_\lambda \end{bmatrix}$$

The adjusted test statistic is:

$$RS^*_\tau = [d_\tau - J_{(\tau\delta.\beta).\lambda} J^{-1}_{(\delta.\beta).\lambda} d_\delta - J_{\tau\lambda.(\delta.\beta)} J^{-1}_{\lambda.(\delta.\beta)} d_\lambda][J_{\tau.\beta} - J_{(\tau\delta.\beta).\lambda} J^{-1}_{(\delta.\beta).\lambda} J_{(\delta\tau.\beta).\lambda} - J_{\tau\lambda.(\delta.\beta)} J^{-1}_{\lambda.(\delta.\beta)} J_{\lambda\tau.(\delta.\beta)}]^{-1}$$
$$[d_\tau - J_{(\tau\delta.\beta).\lambda} J^{-1}_{(\delta.\beta).\lambda} d_\delta - J_{\tau\lambda.(\delta.\beta)} J^{-1}_{\lambda.(\delta.\beta)} d_\lambda]' \sim \chi^2_1(0) \quad (117)$$

Lastly, VI) $H_o^g : \lambda = 0$ in presence of $\gamma, \sigma^2_\mu, \rho, \delta, \tau$.

Here, $\phi = (\gamma, \sigma^2_\mu, \rho, \delta, \tau)$
$d_\psi = d_\lambda$
$d_\phi = (d_\gamma, d_{\sigma_{\mu^2}}, d_\rho, d_\delta, d_\tau)$
$J_{\psi\phi.\omega} = (0, 0, 0, J_{\lambda\delta}, J_{\lambda\tau})$.

$$J_{\phi.\omega} \begin{bmatrix} J_{\gamma.\omega} & J_{\gamma\sigma^2_\mu.\sigma^2_e} & J_{\gamma\rho} & 0 & 0 \\ J_{\sigma^2_\mu\gamma.\sigma^2_e} & J_{\sigma^2_\mu.\sigma^2_e} & J_{\sigma^2_\mu\rho} & 0 & 0 \\ J_{\rho\gamma} & J_{\rho\sigma^2_\mu} & J_\rho & 0 & 0 \\ 0 & 0 & 0 & J_{\delta.\beta} & J_{\delta\tau.\beta} \\ 0 & 0 & 0 & J_{\tau\delta.\beta} & J_{\tau.\beta} \end{bmatrix}$$

The adjusted test statistic is:

$$RS^*_\lambda = [d_\lambda - J_{(\lambda\delta).(\tau.\beta)} J^{-1}_{(\delta.\beta).(\tau.\beta)} d_\delta - J_{\lambda\tau.(\delta.\beta)} J^{-1}_{(\tau.\beta).(\delta.\beta)} d_\tau][J_\lambda - J_{(\lambda\delta).(\tau.\beta)} J^{-1}_{(\delta.\beta)(\tau.\beta)} J_{(\delta\lambda).(\tau.\beta)} - J_{\lambda\tau.\delta} J^{-1}_{(\tau.\beta).(\delta.\beta)} J_{\tau\lambda.(\delta.\beta)}]^{-1}[d_\lambda - J_{(\lambda\delta).(\tau.\beta)} J^{-1}_{(\delta.\beta).(\tau.\beta)} d_\delta - J_{\lambda\tau.(\delta.\beta)} J^{-1}_{(\tau.\beta).(\delta.\beta)} d_\tau]' \sim \chi^2_1(0) \quad (118)$$

8.4 Country Lists and Groups

Here are the list of countries in each group divided based on their initial income in 1961:
Panel A (Poorest): The average of real per-capita income has grown by 2.7% over 35 years.
Bangladesh, Benin, Botswana, Burkina Faso, Central African Republic, Chad, Congo Dem. Rep., Ethiopia, Ghana, India, Indonesia, Kenya, Madagascar, Malawi, Mali, Mozambique, Nepal, Niger, Rwanda, Sierra Leone, Sri Lanka, Tanzania, Uganda, Zimbabwe.
Panel B (Middle): The growth rate of the real per capita income is 5.19% from 1961 - 1995.
Angola, Bolivia, Cameroon, Republic of Congo, Cote d'Ivoire, Dominican Republic, Ecuador, Egypt, Honduras, Malaysia, Mauritania, Mauritius, Morocco, Nigeria, Pakistan, Papua New Guinea, Paraguay, Philippines, Senegal, Syria, Thailand, Tunisia, Zambia.
Panel C (Rich): The average of real per capita income has grown by 4.83 % over 35 years.
Argentina, Brazil, Chile, Colombia, Costa Rica, El Salvador, Finland, Guatemala, Hong Kong, Ireland, Jamaica, Japan, Jordan, Korea, Mexico, Nicaragua, Panama, Peru, Portugal, Singapore, South Africa, Spain, Trinidad and Tobago, Turkey, Uruguay.
Panel D (Richest): The average growth rate of per capita income of this group is 4.37 %.
Australia, Austria, Belgium, Canada, Denmark, France, Greece, Israel, Italy, Netherlands, New Zealand, Norway, Sweden, Switzerland, United Kingdom, United States, Venezuela.

Ranking of Countries according to income in 1995:
Poorest: Zimbabwe, Congo Dem. Rep., Burundi, Ethiopia, Central African Republic, Malawi, Mozambique, Madagascar, Niger, Togo, Rwanda, Burkina Faso, Tanzania, Sierra Leone, Ghana, Uganda, Nepal, Kenya, Bangladesh, Benin, Mali, Mauritania, Cote d'Ivoire, Chad, Senegal.
Middle: Zambia, Cameroon, Nigeria, Papua New Guinea, Republic of Congo, Nicaragua, Philippines, Pakistan, India, Angola, Indonesia, Bolivia, Paraguay, Sri Lanka, Morocco, Honduras, Syria, Thailand, Peru, Egypt, Ecuador, Jordan, Tunisia, Guatemala, El Salvador.
Rich: Brazil, Colombia, Panama, Dominican Republic, Uruguay, Mauritius, Botswana, South Africa, Venezuela, Jamaica, Argentina, Costa Rica, Malaysia, Turkey, Chile, Mexico, Portugal, Trinidad and Tobago, Korea, New Zealand, Spain, Israel, Greece, Japan, Finland.
Richest: United States, Italy, Belgium, Norway, Australia, Canada, France, Hong Kong, Netherlands, Denmark, Switzerland, Israel, Sweden, United Kingdom, Greece, Venezuela, Singapore, Ireland, Norway.

As can be seen from the lists and also evident from Figure 1 and its subsequent discussions, there has been some transitional changes among the groups.

8.5 More Monte Carlo Results

In addition to Table 3 - 4, in Section 7, I provide Tables 5 - 8, in addition to support the good finite sample properties of the proposed tests. The tables are listed in the next page onwards.

Table 7: Estimated Rejection Probabilities with $\delta = \tau = \lambda = 0$. Sample size: $N = 25, T = 10$

ρ	η	γ	RS_{γ}^{*}	RS_{γ}	$RS_{\sigma_\mu^2}^{*}$	$RS_{\sigma_\mu^2}$	RS_{ρ}^{*}	RS_{ρ}
0.0	0.2	0.0	0.066	0.955	0.043	0.099	0.072	0.166
0.1	0.2	0.0	0.030	0.997	0.071	0.111	0.112	0.231
0.2	0.2	0.0	0.075	1.000	0.053	0.205	0.192	0.312
0.3	0.2	0.0	0.042	1.000	0.135	0.214	0.274	0.401
0.4	0.2	0.0	0.054	1.000	0.313	0.358	0.440	0.521
0.5	0.2	0.0	0.034	1.000	0.203	0.189	0.712	0.682
0.0	0.0	0.2	0.107	1.000	0.037	0.107	0.035	0.105
0.1	0.0	0.2	0.305	1.000	0.045	0.123	0.127	0.108
0.2	0.0	0.2	0.293	1.000	0.053	0.253	0.120	0.118
0.3	0.0	0.2	0.331	1.000	0.091	0.147	0.209	0.499
0.4	0.0	0.2	0.401	1.000	0.061	0.311	0.219	0.881
0.5	0.0	0.2	0.554	1.000	0.051	0.556	0.293	0.994
0.0	0.0	0.4	0.611	1.000	0.041	0.031	0.051	0.182
0.1	0.0	0.4	0.602	1.000	0.042	0.082	0.104	0.336
0.2	0.0	0.4	0.742	1.000	0.061	0.141	0.202	0.692
0.3	0.0	0.4	0.739	1.000	0.038	0.231	0.284	0.944
0.4	0.0	0.4	0.779	1.000	0.049	0.419	0.335	0.995
0.5	0.0	0.4	0.891	0.999	0.001	0.627	0.207	1.000
0.0	0.2	0.2	0.105	1.000	0.184	0.207	0.035	0.191
0.1	0.2	0.2	0.204	1.000	0.304	0.313	0.125	0.302
0.2	0.2	0.2	0.305	1.000	0.305	0.545	0.112	0.351
0.3	0.2	0.2	0.312	1.000	0.502	0.790	0.201	0.565
0.4	0.2	0.2	0.441	1.000	0.589	0.811	0.304	0.628
0.5	0.2	0.2	0.618	1.000	0.612	0.812	0.419	0.920
0.2	0.0	0.0	0.082	1.000	0.056	0.159	0.097	0.118
0.2	0.1	0.0	0.077	1.000	0.062	0.312	0.109	0.212
0.2	0.2	0.0	0.093	0.998	0.164	0.302	0.118	0.318
0.2	0.3	0.0	0.092	0.999	0.161	0.316	0.199	0.401
0.2	0.4	0.0	0.075	0.998	0.251	0.515	0.100	0.399
0.2	0.5	0.0	0.094	0.998	0.298	0.532	0.109	0.333
0.0	0.0	0.2	0.107	1.000	0.060	0.117	0.032	0.092
0.0	0.1	0.2	0.118	1.000	0.097	0.105	0.036	0.108
0.0	0.2	0.2	0.204	1.000	0.104	0.209	0.024	0.112
0.0	0.3	0.2	0.204	1.000	0.203	0.306	0.022	0.308
0.0	0.4	0.2	0.307	1.000	0.335	0.413	0.036	0.410
0.0	0.5	0.2	0.406	1.000	0.416	0.508	0.028	0.399

Table 8: Estimated Rejection Probabilities with $\delta = \tau = \lambda = 0$. Sample size: $N = 25, T = 10$

ρ	η	γ	RS_{γ}^{*}	RS_{γ}	$RS_{\sigma_\mu^2}^{*}$	$RS_{\sigma_\mu^2}$	RS_{ρ}^{*}	RS_{ρ}
0.2	0	0	0.036	0.845	0.041	0.103	0.104	0.301
0.2	0	0.1	0.330	0.990	0.031	0.115	0.147	0.306
0.2	0	0.2	0.275	1.000	0.040	0.151	0.114	0.213
0.2	0	0.3	0.201	1.000	0.051	0.184	0.102	0.394
0.2	0	0.4	0.311	1.000	0.037	0.123	0.105	0.682
0.2	0	0.5	0.314	1.000	0.045	0.152	0.143	0.846
0	0.2	0	0.052	0.970	0.162	0.206	0.042	0.112
0	0.2	0.1	0.141	1.000	0.133	0.301	0.073	0.212
0	0.2	0.2	0.206	1.000	0.205	0.335	0.026	0.102
0	0.2	0.3	0.331	1.000	0.201	0.421	0.051	0.211
0	0.2	0.4	0.312	1.000	0.333	0.533	0.059	0.181
0	0.2	0.5	0.415	1.000	0.398	0.546	0.071	0.253
0	0.4	0	0.041	0.990	0.152	0.550	0.058	0.123
0	0.4	0.1	0.142	1.000	0.225	0.601	0.043	0.233
0	0.4	0.2	0.106	1.000	0.105	0.607	0.027	0.341
0	0.4	0.3	0.204	1.000	0.204	0.715	0.043	0.311
0	0.4	0.4	0.301	1.000	0.301	0.681	0.043	0.368
0	0.4	0.5	0.411	1.000	0.312	0.747	0.039	0.221
0.2	0.2	0	0.070	1.000	0.049	0.302	0.100	0.230
0.2	0.2	0.1	0.212	1.000	0.117	0.315	0.139	0.101
0.2	0.2	0.2	0.307	1.000	0.107	0.449	0.216	0.141
0.2	0.2	0.3	0.301	1.000	0.201	0.568	0.206	0.256
0.2	0.2	0.4	0.412	1.000	0.312	0.621	0.301	0.561
0.2	0.2	0.5	0.512	1.000	0.376	0.632	0.303	0.770
0	0	0.4	0.328	1.000	0.041	0.110	0.057	0.033
0	0.1	0.4	0.399	1.000	0.197	0.215	0.034	0.178
0	0.2	0.4	0.458	1.000	0.104	0.311	0.042	0.108
0	0.3	0.4	0.502	1.000	0.233	0.311	0.047	0.271
0	0.4	0.4	0.555	1.000	0.358	0.529	0.034	0.390
0	0.5	0.4	0.419	1.000	0.427	0.599	0.052	0.383
0.2	0	0.2	0.103	1.000	0.030	0.162	0.113	0.115
0.2	0.1	0.2	0.103	1.000	0.103	0.246	0.114	0.176
0.2	0.2	0.2	0.203	1.000	0.103	0.237	0.202	0.243
0.2	0.3	0.2	0.203	1.000	0.193	0.226	0.314	0.330
0.2	0.4	0.2	0.301	1.000	0.301	0.331	0.416	0.419
0.2	0.5	0.2	0.422	1.000	0.402	0.512	0.515	0.613

Table 9: Estimated Rejection Probabilities with $\gamma = \sigma_\mu^2 = \rho = 0$. Sample size: $N = 25, T = 10$

λ	τ	δ	$RS_\delta{}^*$	RS_δ	$RS_\tau{}^*$	RS_τ	$RS_\lambda{}^*$	RS_λ
0	0.2	0	0.021	0.999	0.907	0.991	0.051	0.997
0.1	0.2	0	0.380	0.999	0.971	0.999	0.061	1.000
0.2	0.2	0	0.051	0.996	0.991	1.000	0.166	1.000
0.3	0.2	0	0.470	0.989	0.984	1.000	0.358	1.000
0.4	0.2	0	0.028	0.965	0.961	1.000	0.639	1.000
0.5	0.2	0	0.039	0.914	0.911	1.000	0.859	1.000
0	0	0.2	0.091	1.000	0.261	0.870	0.037	0.950
0.1	0	0.2	0.112	1.000	0.394	0.968	0.122	0.991
0.2	0	0.2	0.263	1.000	0.328	0.997	0.151	0.997
0.3	0	0.2	0.309	1.000	0.479	0.999	0.127	0.998
0.4	0	0.2	0.398	1.000	0.496	1.000	0.288	1.000
0.5	0	0.2	0.402	0.998	0.593	1.000	0.578	1.000
0	0	0.4	0.116	1.000	0.070	0.977	0.052	0.987
0.1	0	0.4	0.127	1.000	0.341	0.998	0.147	1.000
0.2	0	0.4	0.320	1.000	0.418	1.000	0.149	1.000
0.3	0	0.4	0.309	1.000	0.476	1.000	0.264	1.000
0.4	0	0.4	0.419	1.000	0.489	1.000	0.153	1.000
0.5	0	0.4	0.393	1.000	0.598	1.000	0.314	1.000
0	0.2	0.2	0.115	1.000	0.897	1.000	0.070	1.000
0.1	0.2	0.2	0.109	1.000	0.957	1.000	0.116	1.000
0.2	0.2	0.2	0.271	1.000	0.979	1.000	0.243	1.000
0.3	0.2	0.2	0.314	1.000	0.990	1.000	0.199	1.000
0.4	0.2	0.2	0.519	1.000	1.000	1.000	0.241	1.000
0.5	0.2	0.2	0.602	0.999	0.998	1.000	0.485	1.000
0	0	0.2	0.081	1.000	0.086	0.881	0.056	0.950
0	0.1	0.2	0.102	1.000	0.788	0.993	0.069	0.998
0	0.2	0.2	0.363	1.000	0.862	1.000	0.071	1.000
0	0.3	0.2	0.309	1.000	0.923	1.000	0.050	1.000
0	0.4	0.2	0.378	1.000	0.944	1.000	0.080	1.000
0	0.5	0.2	0.410	1.000	0.936	1.000	0.072	1.000
0	0	0.4	0.092	1.000	0.094	0.985	0.058	0.990
0	0.1	0.4	0.202	1.000	0.822	1.000	0.028	0.999
0	0.2	0.4	0.343	1.000	0.885	1.000	0.036	1.000
0	0.3	0.4	0.409	1.000	0.915	1.000	0.041	1.000
0	0.4	0.4	0.418	1.000	0.912	1.000	0.041	1.000
0	0.5	0.4	0.471	1.000	0.932	1.000	0.047	1.000

Table 10: Estimated Rejection Probabilities with $\gamma = \sigma_\mu^2 = \rho = 0$. Sample size: $N = 25, T = 10$

λ	τ	δ	$RS_\delta{}^*$	RS_δ	$RS_\tau{}^*$	RS_τ	$RS_\lambda{}^*$	RS_λ
0	0.2	0	0.021	0.999	0.907	0.991	0.051	0.997
0.1	0.2	0	0.380	0.999	0.971	0.999	0.061	1.000
0.2	0.2	0	0.051	0.996	0.991	1.000	0.166	1.000
0.3	0.2	0	0.470	0.989	0.984	1.000	0.358	1.000
0.4	0.2	0	0.028	0.965	0.961	1.000	0.639	1.000
0.5	0.2	0	0.039	0.914	0.911	1.000	0.859	1.000
0	0	0.2	0.091	1.000	0.261	0.870	0.037	0.950
0.1	0	0.2	0.112	1.000	0.394	0.968	0.122	0.991
0.2	0	0.2	0.263	1.000	0.328	0.997	0.151	0.997
0.3	0	0.2	0.309	1.000	0.479	0.999	0.127	0.998
0.4	0	0.2	0.398	1.000	0.496	1.000	0.288	1.000
0.5	0	0.2	0.402	0.998	0.593	1.000	0.578	1.000
0	0	0.4	0.116	1.000	0.070	0.977	0.052	0.987
0.1	0	0.4	0.127	1.000	0.341	0.998	0.147	1.000
0.2	0	0.4	0.320	1.000	0.418	1.000	0.149	1.000
0.3	0	0.4	0.309	1.000	0.476	1.000	0.264	1.000
0.4	0	0.4	0.419	1.000	0.489	1.000	0.153	1.000
0.5	0	0.4	0.393	1.000	0.598	1.000	0.314	1.000
0	0.2	0.2	0.115	1.000	0.897	1.000	0.070	1.000
0.1	0.2	0.2	0.109	1.000	0.957	1.000	0.116	1.000
0.2	0.2	0.2	0.271	1.000	0.979	1.000	0.243	1.000
0.3	0.2	0.2	0.314	1.000	0.990	1.000	0.199	1.000
0.4	0.2	0.2	0.519	1.000	1.000	1.000	0.241	1.000
0.5	0.2	0.2	0.602	0.999	0.998	1.000	0.485	1.000
0	0	0.2	0.081	1.000	0.086	0.881	0.056	0.950
0	0.1	0.2	0.102	1.000	0.788	0.993	0.069	0.998
0	0.2	0.2	0.363	1.000	0.862	1.000	0.071	1.000
0	0.3	0.2	0.309	1.000	0.923	1.000	0.050	1.000
0	0.4	0.2	0.378	1.000	0.944	1.000	0.080	1.000
0	0.5	0.2	0.410	1.000	0.936	1.000	0.072	1.000
0	0	0.4	0.092	1.000	0.094	0.985	0.058	0.990
0	0.1	0.4	0.202	1.000	0.822	1.000	0.028	0.999
0	0.2	0.4	0.343	1.000	0.885	1.000	0.036	1.000
0	0.3	0.4	0.409	1.000	0.915	1.000	0.041	1.000
0	0.4	0.4	0.418	1.000	0.912	1.000	0.041	1.000
0	0.5	0.4	0.471	1.000	0.932	1.000	0.047	1.000