Testing for a Structural Break in a Spatial Panel Model*

Aparna Dutta†

Economics Department, Boston University

270 Bay State Road, Boston, MA, USA

November 2014

Abstract

This paper considers the problem of testing for a structural break in the spatial lag parameter in a panel model. It proposes a likelihood ratio test of the null hypothesis of no break against the alternative hypothesis of a single break. The limiting distribution of the test is derived under the null when both the number of individual units N and the number of time periods T is large or N is fixed and T is large. The asymptotic critical values for the test staitistic can be obtained analytically. The paper also proposes a break-date estimator that can be employed to determine the location of the break point following evidence against the null hypothesis. I present Monte Carlo evidence to show that the proposed procedure performs well in finite samples. Finally, the paper considers an empirical application of the test on budget spillovers and interdependence in fiscal policy within US states.

Keywords: Panel model, spatial, structural change, budget.

*I am grateful to my advisors Pierre Perron, Zhongjun Qu and Hiroaki Kaido for their guidance, support and encouragement. I would also like to thank Iván Fernández-Val, Jushan Bai, Laurent Pauwels, Qu Feng, Hide Ichimura, Fan Zhuo, Anindya Chakrabarti and seminar participants at Boston University for useful conversations and feedback. I am thankful to Satadru Sengupta for introducing me to Spatial Econometrics. All errors are my own

†apdutta@bu.edu

1 Introduction

Spatial dependence represents a situation where values observed at one location or region depend on the values of neighboring observations at nearby locations. One may ask two questions: first, does this dependence stay the same over time; and second, what might cause the dependence to change? This paper answers the first question by proposing a likelihood ratio test of the null hypothesis of no change against the alternative hypothesis of a one-time change. In case there is evidence against the null hypothesis, the paper consequently proposes a break-date estimator. The second question has been reflected upon through an empirical application of budget spillovers in US states.

In the setup of spatial panel models with N individual units (geographic locations such as countries and zip codes or network units like firms and individuals) observed over T number of periods, where the outcome of each unit depends on its "neighbor's" outcome, there exists a problem of endogeneity. Hence such models are estimated using maximum likelihood or generalized method of moments. Similar to the univariate time series case, in this paper a *sup* LR test is proposed and the asymptotics are derived for large T cases.

In comparison to the vast literature on change point for univariate series, the corresponding literature for panel data is quite small. One of the most popular and early tests in the univariate literature is the popular F test of Chow [1960], which has been modified for cases of unknown and multiple break dates in Andrews [1993], Andrews and Ploberger [1994] and Bai and Perron [1998] among others. Bai [1997], Bai et al. [1998] and Qu and Perron [2007] have extended the single time-series break models to multiple ones. They show that using multiple time series improves estimation precision of the break dates and the size/power of the tests. Perron [2006] provides a survey of the literature.

In the panel data literature, Bai [2010] establishes consistency of the estimated common break point, achievable even if there is a single observation in a regime. The paper proposes a new framework for developing the limiting distribution for the estimated break point and lays down steps to construct confidence intervals. The least squares method is used for estimating breaks in means. Feng et al. [2009] study a multiple regression model in a panel setting where a break occurs at an unknown common date. They establish consistency and rate of convergence both for a fixed time horizon and large panels. In Feng et al. [2009] the limiting distribution is derived without the assumption of shrinking magnitude of break. Liao [2008] uses the Bayesian method for estimation and inference about structural breaks in panel.

Han and Park [1989] develop a multivariate cusum test in order to test for a structural break in panel data and they apply the test on US manufacturing goods trade data. Emerson and Kao [2000] propose two classes of test statistics for detecting a break at an unknown date in panel data models with time trend. The first is a fluctuation test while the second is based on the mean and exponential Wald statistics of Andrews and Ploberger [1994] and maximum Wald statistic of Andrews [1993]. Wachter and Tzavalis [2012] develop a break detecting testing procedure for the AR(p) linear panel data with exogenous or predetermined regressors. The method accommodates structural break in the slope parameters as well as fixed effects and no assumption is imposed on the homogeneity of cross-sectional fixed effects. Pauwels et al. [2012] provide a structural break test for heterogeneous panel data models, where the break affects some but not all cross-section units in the panel. The test is robust to auto-correlated errors. The test statistic is based on comparing pre and post break sample statistics as in Chow [1960].

A higher availability of geocoded socio-economic datasets has led to a vast expansion of the study of spatial interaction between economics agents. Moreover, the recursive relationship between agents in a network can be modeled using spatial econometric methods. Spatial dependence is the transmission of developments across "neighboring" agents. Elhorst [2010] provides detailed methodologies for estimating spatial panels and to test the competing models. The above tests in the panel literature do not explicitly consider the endogeneity problem in the model, which is arises from the spatial dependence. This paper considers a spatial autoregressive model and tests the break in the spatial lag parameter. To test the change in the spatial dependence parameter the paper proposes *sup* LR test similar to Bai [1999]. Yu et al. [2008] and Lee and Yu [2010] provide the asymptotic properties of quasi-maximum likelihood estimators for spatial autoregressive panel data models with fixed effects. The results from Yu et al. [2008] are used to derive the limit distribution of the *sup* LR test in this paper for large T. An estimator for the break date is proposed that can be employed once evidence against no break in the spatial lag parameter is obtained. The performance of this estimator as well as the proposed test statistic in small samples is evaluated via an extensive Monte Carlo study. Wied [2013] develops a CUSUM-type test for time-varying parameters in a spatial autoregressive model for stock returns. The power of the *sup* LR test is compared with the CUSUM test in the paper.

Case et al. [1993] show that a state's budget expenditure depends on the spending of similar[1] states. Therefore, a rise in a "neighboring" state's expenditure results in an increase

[1] Case et al. [1993] define similar states in three different ways - 1) similar in location, 2) similar in income 3) similar in racial composition.

in the state's own expenditure. As an empirical application, this paper applies the likelihood ratio test on the budget dependence of US states over time. The data consists of annual observations for the continental United States during the period 1960-2011. States that are economically similar are defined as neighbors. The test result shows that the null hypothesis of no break in the spatial dependence parameter is rejected and the break date is estimated as 1982. The budget spillover is more pronounced post break. Details of the results and intuitions on why there might be a break are discussed.

The paper is organized as follows: in section 2, the spatial lag model is presented and discussed. Section 3 provides the motivating examples where the test can be applied. The paper proposes a *sup* LR test, which is described in section 4. The limiting distribution of the test is stated in section 5. The outline of the proof is also provided in this section (details are in the Appendix). In the event of rejection of null hypotheis, the paper proposes a break date estimator in section 6. The finite sample properties of the test and the estimator are discussed in section 7. Finally, the paper applies the test in the empirical application of budget spillovers in US states, in section 8. It shows that there was a change in the budget dependence between similar income states.

2 Spatial Lag Model

Let us consider a simple pooled linear regression model

$$y_{it} = x_{it}\beta + \epsilon_{it}, \tag{1}$$

where i is an index of cross-sectional dimension, with i = 1,...,N, and t is an index for the time dimension, with t = 1,...,T. The paper discusses all the results using "time" as the second dimension; however for a general spatial lag model, the second dimension could very well reflect another cross-sectional characteristic, such as the industry sector or the number of classes or groups. y_{it} is an observation on the dependent variable at i and t, x_{it} a $1 \times K$ vector of observations on the (exogenous) explanatory variables including the intercept, β a matching $K \times 1$ vector of regression coefficients, and ϵ_{it} an error term. In stacked form, the simple pooled regression can be written as

$$y = x\beta + \epsilon, \tag{2}$$

with y a $NT \times 1$ vector, X a $NT \times K$ matrix and ϵ a $NT \times 1$ vector. In general, spatial dependence is present whenever the correlation across cross-sectional units is non-zero, and the pattern of non-zero correlations conforms to a specified neighbor relation. When the spatial correlation pertains to the dependent variable it is known as a spatial lag model. The neighbor relation is expressed by means of a spatial weight matrix.

A spatial weights matrix W is a $N \times N$ positive matrix in which the rows and columns correspond to the cross-sectional observations. An element w_{ij} of the matrix expresses the prior strength of the interaction between location i (in the row of the matrix) and location j (column). This can be interpreted as the presence and strength of a link between nodes (the observations) in a network representation that matches the spatial weights structure. In the simplest case, the weights matrix is binary, with $w_{ij} = 1$ when i and j are neighbors, and $w_{ij} = 0$ when they are not. The choice of the weights is typically driven by geographic criteria such as contiguity (sharing a common border) or distance. However, generalizations that incorporate notions of "economic" distance are increasingly being used as well. By convention, the diagonal elements $w_{ii} = 0$. For computational simplicity and to aid the interpretation of the spatial variables, the weights are almost always standardized such that the elements in each row sum to 1, or, $w^s_{ij} = w_{ij}/\sum_j w_{ij}$. Using the subscript to designate the matrix dimension, with W_N as the weights for the cross-sectional dimension, and the observations stacked, the full $NT \times NT$ weights matrix becomes: $W_{NT} = I_T \otimes W_N$, with I_T as an identity matrix of dimension T.

Unlike the time series case, where "neighboring" observations are directly incorporated into a model specification through a shift operator (example $t - 1$), in the spatial literature the neighboring observations are included in the model specification by applying a spatial lag operator (W) to the dependent variable. A spatial lag operator constructs a new variable which consists of the weighted average of the neighboring observations, with the weights as specified in W. The spatial lag model or mixed regressive spatial autoregressive model includes a spatially lagged dependent variable as an explanatory variable in the regression specification. The word "spatial lag" is used to specify the inclusion of the neighboring observations. Similar to time series "lag operator", Wy emphasizes the first-order location lag in the dependent variable. The spatial lag model can be written as

$$y = \rho(I_T \otimes W_N)y + X\beta + \epsilon \tag{3}$$

where ρ is the spatial autoregressive parameter, and the parameter of interest in this paper.

2.1 Endogeneity Problem

The problem in the estimation of the model (3) is that, unlike the time series case, the spatial lag term is endogenous. This is the result of the two-directionality of the neighbor relation in space ("I am my neighbor's neighbor"), in contrast to the one-directionality in time dependence. Rewriting equation (3) in a reduced form:

$$y = [I_T \otimes (I_N - \rho W_N)^{-1}]X\beta + [I_T \otimes (I_N - \rho W_N)^{-1}]\epsilon \tag{4}$$

specifying that the joint determination of the values of the dependent variable in the spatial system is a function of the explanatory variables and error terms at all locations in the system. The presence of the spatially lagged errors in the reduced form illustrates the joint dependence of $W_N y_t$ and ϵ_t in each cross-section. In model estimation, the simultaneity is usually accounted for through instrumentation (IV and GMM estimation) or by specifying a complete distributional model (maximum likelihood estimation). In this paper, I use maximum likelihood estimation.

2.2 Maximum Likelihood Estimation

Assuming Gaussian distribution for the error term, with $\epsilon \sim N(0, \sigma_\epsilon^2 I_{NT})$, the log-likelihood can be written as:

$$lnL = -\frac{NT}{2} ln 2\pi\sigma_\epsilon^2 + Tln|I_N - \rho W_N| - \frac{1}{2\sigma_\epsilon^2}\epsilon'\epsilon \tag{5}$$

where $\epsilon = y - \rho(I_T \otimes W_N)y - X\beta$ and $|I_T \otimes (I_N - \rho W_N)| = Tln|I_N - \rho W_N|$ is the Jacobian of the spatial transformation. To avoid singularity or explosive processes, the parameter space P for the true spatial autoregressive parameter ρ is compact and ρ_0 is in the interior of P.

Lee [2004] discusses the asymptotic properties of the maximum likelihood estimators for the cross-section case. Lee and Yu [2010] and Yu et al. [2008] derive the properties for the spatial panel model with fixed effects. This paper uses the properties of the maximum likelihood estimators to derive the asymptotic distribution of the test statistic.

3 Motivation

The paper considers the following problem in a spatial lag model:

$$y_{it} = \begin{cases} x_{it}\beta + \rho_1 \sum_{j=1}^{N} w_{ij}y_{jt} + \epsilon_{it} & \text{for } t = 1, ..., k_o, \\ x_{it}\beta + \rho_2 \sum_{j=1}^{N} w_{ij}y_{jt} + \epsilon_{it} & \text{for } t = k_o + 1, ..., T \end{cases} \tag{6}$$

$\rho_1 \neq \rho_2$ means there is a change at an unknown date k_0. The paper proposes a *sup* LR test of the null hypothesis of $\rho_1 = \rho_2$ against the alternative hypothesis of a change: $\rho_1 \neq \rho_2$. The test detects structural break in the spatial dependence parameter. Following are some empirical models where the test can be applied, providing motivation for the test.

3.1 Sectoral Output

Acemoglu et al. [2012] look into the intersectoral input-output linkages in the US and show how microeconomic idiosyncratic fluctuations lead to aggregate fluctuations. Defining the sectoral production function as,

$$x_i = z_i l_i^{\alpha} \prod_{j=1}^{n} x_{ij}^{\beta w_{ij}} \tag{7}$$

where x_i is the output of sector i, l_i is the amount of labor hired by the sector, $\alpha \in$ (0,1) is the share of labor, x_{ij} is the amount of commodity j used in the production of good i, and z_i is the idiosyncratic productivity shock to sector i. The exponent $w_{ij} \geq 0$ designates the share of good j in the total intermediate input use of firms in sector i. In particular, w_{ij} = 0 if sector i does not use good j as input for production.

Acemoglu et al. [2012] assume that the input shares of all sectors add up to 1, so $\sum_j w_{ij}$ = 1. With the assumption of market clearing, equation (7) can be rewritten (taking log on both sides) as equation (3). In this case, labor will be an exogenous variable and $\beta_1 \neq \beta_2$ would mean changes in the Cobb-Douglas parameter over time.

3.2 Cigarette Sales

Baltagi and Li [2004] estimate a demand model for cigarettes based on a panel from 46 US states and defining W based on the neighboring states:

$$log(C_{it}) = \beta_1 log(P_{it}) + \beta_2 log(Y_{it}) + \rho \sum_{j=1}^{N} w_{ij} log(C_{jt}) + \epsilon_{it} \tag{8}$$

where C_{it} is real per capita sales of cigarettes by persons of smoking age (14 years and older). This is measured in packs of cigarettes per capita. P_{it} is the average retail price of a pack of cigarettes measured in real terms. Y_{it} is real per capita disposable income. The spatial autocorrelation parameter shows the dependence of cigarette sales in the neighboring states. The tax policy on per packet cigarette differs by states and this leads to substantial cross-state sales. However, over time, tax per packet has become more homogeneous and hence one could expect the parameter ρ to change over time. By testing the hypothesis that $\rho_1 = \rho_2$ against the alternative hypothesis of $\rho_1 \neq \rho_2$, we can check if the dependence on neighboring states has changed over time.

3.3 Budget Spillovers

Case et al. [1993] showed that US states' budget expenditure depends on the spending of similar states:

$$G_{it} = X_{it}\beta + \rho \sum_{j=1}^{N} w_{ij} G_{jt} + \epsilon_{it} \tag{9}$$

where G_{it} is the per capita real government expenditure of state i in year t, X_{it} includes relevant control variables-income and demographic, $w_{ij} > 0$ if a state is the "neighbor" of another state. Case et al. [1993] define "neighbor" in three different ways in their paper - 1) neighbors in location, 2) states having similar income and 3) states having similar racial composition. They found that if the neighboring state increases its budget spending by a dollar, then the state increases its budget expenditure by 70 cents. Policies have changed over the years and one might be interested in testing if the spillover effect remains the same.

3.4 Other Network Motivations

In many of the network studies, the impact of the network is usually estimated by including WY in the model, where W is the weighting matrix defining the network and y is the variable of concern. For example, a weighted average of the math test scores of students sitting beside student i determines student i's test score.

With increasing network data availability, we could have repeated samples from such network experiments and then be curious to know how the impact of the network changes over time. Our structural break test could be used in this respect.

4 Test

In this section, I describe the test statistic. The spatial lag model is given by:

$$y_{it} = x_{it}\beta_t + \rho_t \sum_{j=1}^{N} w_{ij} y_{jt} + \epsilon_{it} \tag{10}$$

where $\epsilon_{it} \sim N(0, \sigma^2_{\epsilon it})$. I want to test the null hypothesis:
$H_0 : \rho_1 = ... = \rho_T$ and $\beta_1 = ... = \beta_T$ and $\sigma^2_{\epsilon i1} = ... = \sigma^2_{\epsilon iT}$
against the alternative
$H_1 : \beta_1 = ... = \beta_T$ and $\sigma^2_{\epsilon i1} = ... = \sigma^2_{\epsilon iT}$ but there is an integer k_0, $1 < k_0 < T$,
such that $\rho_1 = ... = \rho_{k_0} \neq \rho_{k_0+1} = ... = \rho_T$.

Rewriting the panel model with a change point at k_0 in the parameter ρ,

$$y_{it} = \begin{cases} x_{it}\beta + \rho_1 \sum_{j=1}^{N} w_{ij} y_{jt} + \epsilon_{it} & \text{for } t = 1, ..., k_o, \\ x_{it}\beta + \rho_2 \sum_{j=1}^{N} w_{ij} y_{jt} + \epsilon_{it} & \text{for } t = k_o + 1, ..., T \end{cases} \tag{11}$$

$\rho_1 \neq \rho_2$ means there is a change at an unknown date k_0. The problem can be described as testing $\rho_1 = \rho_2$ against $\rho_1 \neq \rho_2$.

Let us write twice the likelihood ratio as

$$2\Lambda_k = 2(lnL_k(\hat{\rho}_k, \hat{\beta}_k, \hat{\sigma}^2_k) + lnL^*_k(\hat{\rho}^*_k, \hat{\beta}_k, \hat{\sigma}^2_k) - lnL_T(\hat{\rho}_T, \hat{\beta}_T, \hat{\sigma}^2_T)), \tag{12}$$

where

- $lnL_k(\hat{\rho}_k, \hat{\beta}_k, \hat{\sigma}^2_k)$ is the log-likelihood defined for t = 1,..,k
- $lnL^*_k(\hat{\rho}^*_k, \hat{\beta}_k, \hat{\sigma}^2_k)$ is the log-likelihood defined for t = k+1,...,T
- $lnL_T(\hat{\rho}_T, \hat{\beta}_T, \hat{\sigma}^2_T)$ is the log-likelihood defined for t =1,...,T

As k_0 is unknown, I use a maximally selected likelihood ratio and reject H_0 if

$$Z_t = \max_{1<k<T} 2\Lambda_k \tag{13}$$

is large. So the suggested test mechanism is to calculate the difference between the log-likelihood under an alternative hypothesis and the log-likelihood under null for every $1 < k < T$ and then the test statistic is the maximum difference between them.

5 Limiting Distribution

In this section I derive the asymptotic distribution of the test statistic. However, before that I specify the assumptions.

5.1 Assumptions

Assumptions on W_N:

Assumption 1. *$w_{ij} \geq 0$, $i \neq j$ for the off-diagonal elements of the spatial weight matrix W_N and its diagonal elements satisfy $w_{n,ii} = 0$ for i = 1,..,N.*

Assumption 2. *W_N is uniformly bounded in both row and column sums.*

Assumption 3. *$|I_T \otimes (I_N - \rho W_N)|$ is invertible for all $\rho \in \mathrm{P}$; moreover, P is compact and ρ_0 is in the interior of P.*

Assumptions on X and ϵ:

Assumption 4. *ϵ_{it} are iid across i and t with $\epsilon \sim N(0, \sigma_\epsilon^2 I_{NT})$ and $E|\epsilon_{it}|^{4+\eta} < \infty$ for some $\eta > 0$.*

Assumption 5. *The matrices $\frac{1}{Nj}\sum_{i=1}^{N}\sum_{t=1}^{j} X_{it}X_{it}'$ and $\frac{1}{Nj}\sum_{i=1}^{N}\sum_{t=j+1}^{T} X_{it}X_{it}'$ have minimum eigen values bounded away from zero in probability for large j or both large N and j. Also, it is assumed that $E||X_{it}^4|| < \infty$.*

Assumption on N and T:

Assumption 6. *N is a non-decreasing function of T and $T \to \infty$*

Assumption 1 is a standard normalization assumption in spatial econometrics while Assumption 2 is also used in Lee [2004] and Yu et al. [2008]. Assumption 3 guarantees that model (4) is valid. Also, compactness is a condition for theoretical analysis. In empirical application, where W_N is row-normalized, one just searches over (-1,1). Assumption 4 provides regularity assumption for ϵ_{it}. Assumption 5 makes sure that the regressors are asymptotically stationary. Assumption 6 allows two cases: (i) $N \to \infty$ as $T \to \infty$ such that $\frac{N}{T} \to k < \infty$, for $k > 0$ and (ii) N is fixed as $T \to \infty$.

Theorem. *Let $\Longrightarrow$ denote weak convergence in distribution under the Skorohod topology. Under assumptions 1-6 and H_0, the limiting distribution of Z_t is*

$$Z_t \Longrightarrow \sup_{s \in (u, 1-u)} \frac{B_1^2(s)}{s(1-s)} \tag{14}$$

where $B_1(s)$, is a standard Brownian bridge and u is a small positive number.
For a known break k_0

$$Z_t \xrightarrow{D} \chi^2(1) \tag{15}$$

To prove the result, I first take a Taylor approximation of $2\Lambda_k$ around the true parameter ρ_0. It is found that the approximations involve partial sums of Gaussian random vectors that are independently and identically distributed. Using results from the maximum likelihood estimation of the Spatial panel model I obtain uniform convergence to Weiner processes. As a next step, the partials sums are manipulated to obtain a Brownian bridge distribution. For a fixed k, it is then easy to show that the asymptotic distribution is Chi-Square. The detailed proof is provided in the Appendix.

The intuition as to why the asymptotic distribution from the univariate time series test is still valid in this case - because the spatial dependence is contained in time, the dependent variable of unit i only depends on the contemporaneous dependent variable of the neighboring units. So the endogeneity does not spread over time and hence the distribution is similar to the one found in univariate time series case.

There is an explicit form of the distribution function of the limit random variable. The critical values are provided in Kiefer [1959], p.438. Some of the relevant critical values are for size = 10%, 1.4978; for size = 5%, 1.8444 and for size = 1% it is 2.649. In this paper I use 5% trimming, i.e. $u = 0.05$.

6 Estimation

Following the evidence against null hypothesis, it is important to determine the location of the break date. The proposed estimator of the break date is the one that maximizes the likelihood under the alternative hypothesis,

$$k_0 = arg\max_k \, lnL_A \tag{16}$$

where lnL_A is the log likelihood under the alternative defined as: $lnL_A = lnL_k + lnL_k^*$

$$lnL_k = -\frac{Nk}{2} ln2\pi\sigma_\epsilon^2 + kln|I_N - \rho W_N| - \frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{N}\sum_{t=1}^{k}\epsilon_{it}\epsilon_{it}$$

$$lnL_k^* = -\frac{N(T-k)}{2} ln2\pi\sigma_\epsilon^2 + (T-k)ln|I_N - \rho W_N| - \frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{N}\sum_{t=k+1}^{T}\epsilon_{it}\epsilon_{it}$$

where lnL_k is the log-likelihood defined for t = 1,..,k and lnL_k^* is the log-likelihood defined for t = k+1,...,T

The asymptotic properties of the estimator, including the consistency, rate of convergence, and limit distribution are currently under investigation. Simulation evidence, presented is section 7, shows that the estimator performs very well in small samples in terms of bias and root mean squared error. The root mean squared error is shown to decrease as the sample size increases, thereby suggesting that the estimator is indeed consistent.

7 Monte Carlo Results

To evaluate the finite sample performance of the LR test and the performance of the estimator, this section reports results of a limited set of sampling experiments. All results reported are for 1000 simulations. I consider the data generating process -

$$y_{it} = \begin{cases} 1 + x_{it} + 0.6 \sum_{j=1}^{N} w_{ij} y_{jt} + \epsilon_{it} & \text{pre-break} \\ 1 + x_{it} + \rho_2 \sum_{j=1}^{N} w_{ij} y_{jt} + \epsilon_{it} & \text{post-break} \end{cases} \tag{17}$$

where x_{it} from $N(0,1)$ and ϵ_{it} from $N(0,1.3)$

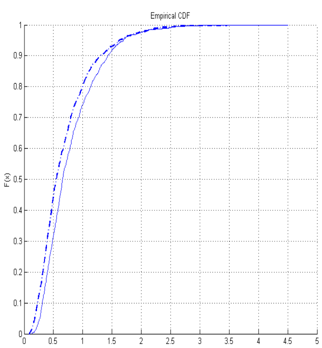
I first look into the power of the proposed test. Let $\rho_1 = 0.6$ and the actual break date is $k_0 = T/2$ in each of the cases. I find that the test has high power even with N and T = 50 as seen in Table 1. The power increases with increases in N and/ or T (in Table 2).

Table 1: Power of the test- I

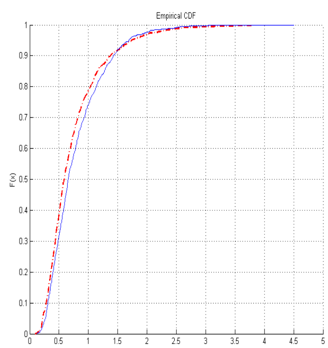
N	T	Rho2	Frequency of rejection
50	50	0.7	0.957
50	50	0.65	0.337
50	50	0.55	0.263
50	50	0.5	0.807
50	50	-0.6	1

Table 2: Power of the test - II

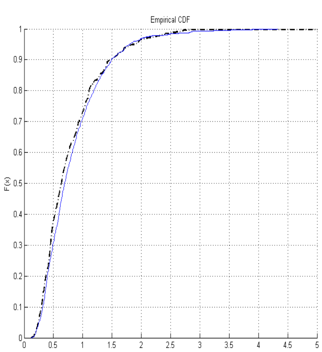
N	T	Rho2	Frequency of rejection
50	100	0.65	0.657
50	100	0.55	0.551
50	200	0.65	0.932
50	200	0.55	0.881
100	50	0.65	0.515
100	50	0.55	0.401
100	100	0.65	0.852
100	100	0.55	0.741
100	200	0.65	0.989
100	200	0.55	0.971



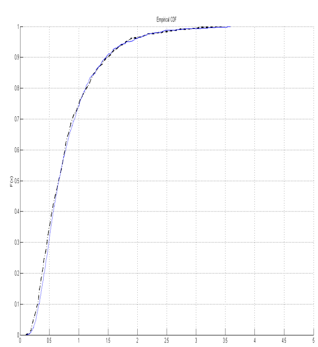
(a) N = 50, T= 50



(b) N = 200, T = 200



(c) N = 50, T = 500



(d) N = 500, T= 500

Figure 1: Emprical versus Asymptotic Distribution

Next I look into graphical comparisons between empirical and asymptotic distributions presented in Figure 1. The continuous lines are the asymptotic distributions and the dotted lines are the empirical cdf. It is found, that even with a small T, there is no size distortion and the empirical distribution matches closely the asymptotic distribution. As T increases, the two distributions overlap.

For a known break, the asymptotic distribution is chi-square with 1 degree of freedom. The graphical comparison presented in Figure 2 shows that even with N= 50, T= 50, with a known break, the empirical distribution is very close to the asymptotic chi-square distribution.

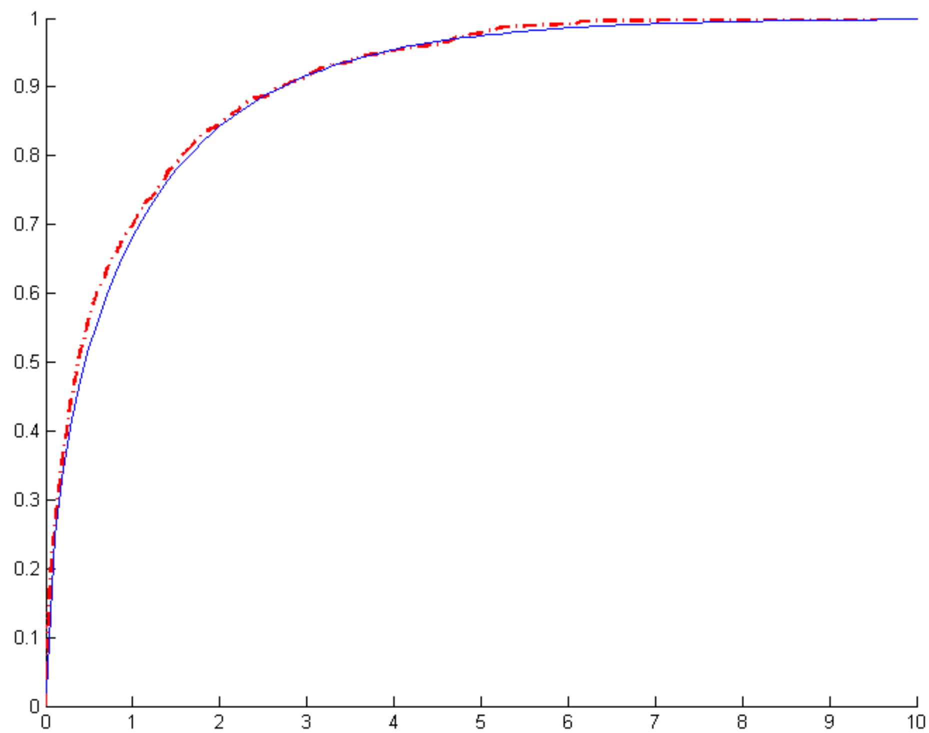


Figure 2: CDF plot for empirical distribution with a known break

Next I compare the performance of the break-date estimator. The bias is almost negligible. The root mean square decreases with increases in N. With increases in T the standard deviation does not go down. This is a well known result in the univariate time series literature-only the break fraction can be consistently estimated, not the break date.

Also, I make a quick comparison with the ordinary least squares residuals-based method, with the estimator defined by

$$\hat{k} = arg \min_{1 \le k \le T} SSR(k) \tag{18}$$

Here $SSR(k)$ is the sum of squared residuals of the model under the alternative assuming a break at date k. The bias is comparable in the two cases, but the standard deviation and root mean square is higher for the OLS residual-based estimate of break date.

Table 3: Estimator Performance - Likelihood Method

Rho1	Rho2	N	T	Break date	bias	St.dev	RMSE
0.6	0.7	50	50	25	0.1	1.01	1.01
0.6	0.7	50	100	50	0.08	1.16	1.16
0.6	0.7	50	200	100	0.11	1.1	1.1
0.6	0.7	50	50	25	0.1	1.01	1.01
0.6	0.7	100	50	25	0.04	0.67	0.67
0.6	0.7	200	50	25	0.01	0.23	0.23
0.6	0.7	50	50	25	0.1	1.01	1.01
0.6	0.7	100	100	50	0.06	0.52	0.53
0.6	0.65	50	50	25	0.35	5.77	5.78
0.6	0.55	50	50	25	0.16	6.99	6.99
0.6	-0.6	50	50	25	0	0	0

Looking at the tables closely, an interesting pattern is observed, there is an asymmetry in the behavior of the estimator and the power of the test. When $\rho_2 = 0.55$ the power of the test is poorer compared to that when $\rho_2 = 0.65$. Similarly the break date estimator has a lower standard deviation and root mean square when the post-break parameter is increasing ($\rho_2 = 0.65$) as compared to a comparable reduction in the post-break parameter ($\rho_2 = 0.55$). An explanation for such behavior could be that, when post-break parameter is increasing ($\rho_2 = 0.65$), there is a higher signal of spatial dependence. This leads to reduction in the variance and makes it easier to determine the break. However when the post-break parameter is comparably lower ($\rho_2 = 0.55$) the signal is lower giving rise to more variation and difficult to capture the break.

Table 4: Estimator Performance - OLS Residuals

rho1	rho2	N	T	Break date	bias	St.dev	RMSE
0.6	0.7	50	50	25	-0.2	2.53	2.54
0.6	0.7	50	100	50	-0.31	2.01	2.03
0.6	0.7	50	200	100	-0.36	1.85	1.88
0.6	0.7	50	50	25	-0.2	2.53	2.54
0.6	0.7	100	50	25	-0.14	1.17	1.18
0.6	0.7	200	50	25	-0.09	0.49	0.5
0.6	0.7	50	50	25	-0.2	2.53	2.54
0.6	0.7	100	100	50	-0.22	1.09	1.11
0.6	0.65	50	50	25	-0.51	8.95	8.96
0.6	0.55	50	50	25	-0.03	9.8	9.8
0.6	-0.6	50	50	25	0	0	0

The proposed likelihood based estimator performs well in finite sample. As N increases, the root mean square error decreases suggesting that the estimator is consistent.

8 Budget Spillovers

Case et al. [1993] showed how a US state's budget expenditure depends on the spending of similar states. Quoting Arkansas state Senator Doug Brandon (1989) describing his state's budgetary policy as

> "We do everything everyone else does."

The proposed *sup* LR test is used to check the hypothesis that a state's dependence on another's budget remained the same in the US or has changed over time. The data consists of an annual panel of US states from 1960 to 2011. All dollar figures are calculated on a per capita basis and deflated using the GDP deflator (the base year being 2009). The dependent variable is the government expenditure of state i in the year t. The budget expenditure is the sum of the direct spending of state and local governments. The variables included in X_{it} other than the intercept are: the real per-capita personal income, income squared, real per capita total intergovernmental federal revenue to state and local governments, population density, proportion of the population at least 65 years old, proportion of the population between 5 and 14 years old, and proportion of the population that is black. The income and revenue are the resources the state government can use. The square of the income picks up possible non-linear effects of changing resources. The population density captures the possibility that there are potential congestion effects and scale economies in the provision of state and local government services. States with different age and racial structures may have different demands for publicly provided goods. Hence demographic variables are included.

The model can be written as:

$$G_{it} = X_{it}\beta + \rho \sum_{j=1}^{N} w_{ij} G_{jt} + \epsilon_{it} \tag{19}$$

where X includes all the control variables. I consider T= 52 from 1960 to 2011 and N = 49 states in the US. Case et al. [1993] use three different ways to define the weight matrix. In this paper, I define it as $w_{ij} = (1/|Y_i - Y_j|)/S_i$, where Y_k is the mean income over the sample period and S_i is the sum $\sum_j 1/|Y_i - Y_j|$. According to this definition of the weight matrix, rich states are neighbors to rich states and poor states are neighbors to poor states.

Table 5: Full model estimate

	Coefficient	Asymptotic t-stat	p-value
intercept	0.697432	0.214295	0.830317
Pop65	-0.404178	-4.898872	0.000001
Pop5to14	-0.058942	-0.57394	0.566009
Popblack	-0.056243	-4.30405	0.000017
Popden	-0.000282	-2.213868	0.026838
F	1.735213	58.255465	0
Y	0.130133	14.288976	0
Y^2	0.000018	1.621974	0.104809
$W * G$	0.121999	7.302404	0

All the test results are based on tests with size 5%. We reject the null hypothesis of no break, implying evidence for a break. The break date is estimated at 1982. The pre-break budget spillover coefficient is estimated as 0.0229 while the post-break budget spillover coefficient is estimated as 0.1056. As to why there might be a break, there could be two reasons: 1) In 1981, Ronald Reagan became the president of the United States and advocated many different policies across US states (also known as Reagonomics). 2) The number of Democratic governors in the US started decreasing post 1983 suggesting synchronized Republican economic policies in different states.

To differentiate between trend behaviors and fluctuations, a Hodrick-Prescott filter is applied on all the dollar value variables to closely look into idiosyncratic budget spillovers in US states. We reject the null hypothesis of no break. The break date is then estimated to be in 1977. The pre-break ρ coefficient is 0.5718 and the post-break ρ coefficient is 0.3746. Firstly this suggests that the idiosyncrasy in budget expenditure for a state depends on "similarly" situated states. Secondly, the dependence goes down post-break. This can be attributed to more power given to the governors in the 1980's. For the federal government (central planner) the budget policies for each state will be similar; compared to individual governors in each state who will adjust the budget expenditures for their states based on individual needs. So overall even though the spillovers increase (capturing overall trend in the economy), the budget spillovers in case of idiosyncracies reduce over time.

9 Conclusion

This paper considers the problem of structural break in the spatial dependence parameter in a panel model and provides a likelihood ratio test.

In this paper, I first describe the spatial panel model and the interpretation of the spatial lag or spatial autoregressive parameter. Next I motivate the problem of structural break in such parameter. The *sup* LR test statistic is proposed and under large T, the limiting distribution is derived. The test is easy to implement and the critical values can be analytically obtained.

In case there is evidence to reject the null hypothesis, the paper proposes a break date estimator based on the argument which maximizes the likelihood ratio. The finite sample properties of the test and the break-date estimator are provided. The monte carlo simulations show that the test has good power even in small samples. The estimator of the break date shows negligible bias and the root mean square decreases with increases in N suggesting a consistent break-date estimator for a panel model.

The paper then considers the problem of budget spillovers across US states and the change in the spatial dependence over time. The test rejects the null hypothesis of no break in budget spillovers for 1) the spillover in the overall budget expenditure of US states and 2) the spillover in the fluctuations of budget expenditure. The overall trend of spatial dependence in budget expenditure is found to have increased post-break, but the idiosyncrasies in budget expenditure are less spatially dependent post-break.

The following extensions to the paper are being considered - 1) asymptotic limit distribution of the test staitstic for large N, (2) proving the consistency of the break date estimator and deriving the limiting distribution, and 3) extending the test to multiple structural breaks.

10 Appendix: Proof of theorem

Let $\theta = (\rho, \beta, \sigma_\epsilon^2)$. Then,

$$
\begin{aligned}
lnL_T(\theta) &= -\frac{NT}{2}ln2\pi\sigma_\epsilon^2 + Tln|I_N - \rho W_N| - \frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{N}\sum_{t=1}^{T}\epsilon_{it}'\epsilon_{it} \\
lnL_k(\theta) &= -\frac{Nk}{2}ln2\pi\sigma_\epsilon^2 + kln|I_N - \rho W_N| - \frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{N}\sum_{t=1}^{k}\epsilon_{it}'\epsilon_{it} \\
lnL_k^*(\theta) &= -\frac{N(T-k)}{2}ln2\pi\sigma_\epsilon^2 + (T-k)ln|I_N - \rho W_N| - \frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{N}\sum_{t=k+1}^{T}\epsilon_{it}'\epsilon_{it}
\end{aligned}
$$

Denoting $lnL_T(\theta) = L_c$, $lnL_k(\theta) = L_1$ and $lnL_k^*(\theta) = L_2$. Also defining $\hat{\rho}_k$ as the MLE estimate for pre-break regime under alternative, $\hat{\rho}_k^*$ as the MLE estimate for post-break regime under alternative and $\hat{\rho}_T$ as the MLE estimate under null. Taking Taylor expansion of $2[L_1 + L_2 - L_c]$ around the true value ρ_0 and denoting that by R_k

$$
\begin{aligned}
R_k = 2[&L_1(\rho_0) + L_2(\rho_0) - L_c(\rho_0) \\
&+ L_1'(\rho_0)(\hat{\rho}_k - \rho_0) + \frac{L_1''(\rho_0)}{2}(\hat{\rho}_k - \rho_0)^2 \\
&+ L_2'(\rho_0)(\hat{\rho}_k^* - \rho_0) + \frac{L_2''(\rho_0)}{2}(\hat{\rho}_k^* - \rho_0)^2 \\
&- L_c'(\rho_0)(\hat{\rho}_T - \rho_0) + \frac{L_c''(\rho_0)}{2}(\hat{\rho}_T - \rho_0)^2]
\end{aligned}
$$

Now, $L_1(\rho_0) + L_2(\rho_0) = L_c(\rho_0)$. So R_k can be rewritten as:

$$
\begin{aligned}
R_k = [&2L_1'(\rho_0)(\hat{\rho}_k - \rho_0) + L_1''(\rho_0)(\hat{\rho}_k - \rho_0)^2 \\
&+ 2L_2'(\rho_0)(\hat{\rho}_k^* - \rho_0) + L_2''(\rho_0)(\hat{\rho}_k^* - \rho_0)^2 \\
&- 2L_c'(\rho_0)(\hat{\rho}_T - \rho_0) + L_c''(\rho_0)(\hat{\rho}_T - \rho_0)^2]
\end{aligned}
$$

From Lee [2004] and Yu et al. [2008] under the assumptions 1-6

$$\sqrt{NT}(\hat{\rho}_T - \rho_0) = \big[-\frac{1}{NT}L_c''(\rho_0)\big]^{-1}\frac{1}{\sqrt{NT}}L_c'(\rho_0)$$
$$\sqrt{Nk}(\hat{\rho}_k - \rho_0) = \big[-\frac{1}{Nk}L_1''(\rho_0)\big]^{-1}\frac{1}{\sqrt{Nk}}L_1'(\rho_0)$$
$$\sqrt{N(T-k)}(\hat{\rho}_k^* - \rho_0) = \big[-\frac{1}{N(T-k)}L_2''(\rho_0)\big]^{-1}\frac{1}{\sqrt{N(T-k)}}L_2'(\rho_0)$$

Using these relationships and rearranging the terms, R_k can be rewritten as:

$$\begin{aligned} R_k = &\frac{1}{\sqrt{Nk}}L_1'(\rho_0)\big[-\frac{1}{Nk}L_1''(\rho_0)\big]^{-1}\frac{1}{\sqrt{Nk}}L_1'(\rho_0) \\ &+ \frac{1}{\sqrt{N(T-k)}}L_2'(\rho_0)\big[-\frac{1}{N(T-k)}L_2''(\rho_0)\big]^{-1}\frac{1}{\sqrt{N(T-k)}}L_2'(\rho_0) \\ &- \frac{1}{\sqrt{NT}}L_c'(\rho_0)\big[-\frac{1}{NT}L_c''(\rho_0)\big]^{-1}\frac{1}{\sqrt{NT}}L_c'(\rho_0) \end{aligned}$$

Let $G_N = W_N[I_N - \rho_N W_N]^{-1}$ then

$$-\frac{1}{NT}L_c''(\rho_0) = \frac{1}{\sigma_{\epsilon 0}^2}\sum_{t=1}^{T}\big((W_N Y_{Nt})W_N Y_{Nt} + tr(G_N^2)\big)$$

where $W_N Y_{Nt} = G_N X_{Nt}\beta_0 + G_N \epsilon_{Nt}$ and $\underset{T\to\infty}{lim} - \frac{1}{NT}L_c''(\bar{\rho}) \Rightarrow H^{*'}H^*$. Also,

$$\frac{1}{\sqrt{NT}}L_c'(\rho_0) = \frac{1}{\sigma_{\epsilon 0}^2\sqrt{NT}}\sum_{t=1}^{T}\big[(G_N X_{Nt}\beta_0)'\epsilon_{Nt}\big] + \frac{1}{\sigma_{\epsilon 0}^2\sqrt{NT}}\sum_{t=1}^{T}\big[\epsilon_{Nt}'G_N'\epsilon_{Nt} - \sigma_{\epsilon 0}^2 trG_N\big]$$

$$\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\big[\epsilon_{Nt}'G_N'\epsilon_{Nt} - \sigma_{\epsilon 0}^2 trG_N\big] = o_p(1)$$
$$\frac{1}{\sqrt{NT}}\sum_{t=1}^{T}\big[(G_N X_{Nt}\beta_0)'\epsilon_{Nt}\big] = O_p(1)$$

Now, $\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\big[\frac{1}{\sqrt{N}}(G_N X_{Nt}\beta_0)'\epsilon_{Nt}\big] = T^{-1/2}\sum_{t=1}^{T}H_{Nt}\epsilon_{Nt}$. Here, $\frac{1}{\sqrt{N}}(G_N X_{Nt}\beta_0)' = H_{Nt}$ and $\underset{T\to\infty}{lim} H_{Nt} \Rightarrow H^*$. As long as $T\to\infty$, by FCLT we get:

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}H_{Nt}\epsilon_{Nt} \Rightarrow H^* W(1)$$

where $W(t)$ is a standard Weiner process. Thus, if we let $\frac{k}{T} = \lambda$. Then by FCLT,

$$\frac{1}{\sqrt{k}} \sum_{t=1}^{k} H_{Nt} \epsilon_{Nt} \Rightarrow \frac{H^* W(\lambda)}{\sqrt{\lambda}}$$

$$\frac{1}{\sqrt{T-k}} \sum_{t=k+1}^{T} H_{Nt} \epsilon_{Nt} \Rightarrow \frac{H^* (W(1) - W(\lambda))}{\sqrt{1-\lambda}}$$

Hence we get:

$$\begin{aligned} R(k) \Rightarrow & \frac{H^* W(\lambda)(H^*)^{-1}}{\sqrt{\lambda}} \frac{H^* W(\lambda)(H^*)^{-1}}{\sqrt{\lambda}} \\ & + \frac{H^*(W(1) - W(\lambda))(H^*)^{-1}}{\sqrt{1-\lambda}} \frac{H^*(W(1) - W(\lambda))(H^*)^{-1}}{\sqrt{1-\lambda}} \\ & - H^* W(1)(H^*)^{-1} H^* W(1)(H^*)^{-1} \end{aligned}$$

Let

$$R(\lambda) \equiv \frac{1}{\lambda}[W(\lambda)]^2 + \frac{1}{1-\lambda}[W(1) - W(\lambda)]^2 - [W(1)]^2$$

Rearranging the terms we get:

$$\begin{aligned} \sup_{\lambda \in (u, 1-u)} R(\lambda) &\Rightarrow \sup_{\lambda \in (u, 1-u)} \frac{[\lambda W(1) - W(\lambda)]^2}{\lambda(1-\lambda)} \\ \text{or} \sup_{\lambda \in (u, 1-u)} R(\lambda) &\Rightarrow \sup_{\lambda \in (u, 1-u)} \frac{B_1^2(\lambda)}{\lambda(1-\lambda)} \end{aligned}$$

For known k_0, $\lambda_0 = \frac{k_0}{T}$, the limit distribution of $R(\lambda_0)$ is χ_1^2.

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., and Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, 80:1977–2016.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61:821–856.

Andrews, D. W. K. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica*, 62:1383–1414.

Bai, J. (1997). Estimation of a change point in multiple regressions models. *The Review of Economics and Statistics*, 79:551–563.

Bai, J. (1999). Likelihood ratio tests for multiple structural changes. *Journal of Econometrics*, 91:299–323.

Bai, J. (2010). Common breaks in means and variances for panel data. *Journal of Econometrics*, 157:78–92.

Bai, J., Lumsdaine, R., and Stock, J. (1998). Testing and dating common breaks in multivariate time series. *Review of Economic Studies*, 65:395–432.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66:47–78.

Baltagi, B. H. and Li, D. (2004). Prediction in the panel data model with spatial correlation. In Anselin, L., Florax, R., and Rey, S., editors, *Advances in spatial econometrics*, pages 283–295. Springer Berlin Heidelberg.

Case, A. C., Rosen, H. S., and Jr, J. R. H. (1993). Budget spillovers and fiscal policy interdependence: Evidence from the states. *Journal of public economics*, 52.3:285–307.

Chow, G. C. (1960). Test of equality between sets of coefficients in two linear regressions. *Econometrica*, 28:591–605.

Elhorst, J. (2010). Spatial panel data models. In Fischer, M. M. and Getis, A., editors, *Handbook of Applied Spatial Analysis*, pages 377–407. Springer Berlin Heidelberg.

Emerson, J. and Kao, C. (2000). Testing for structural change of a time trend regression in panel data. *Syracuse University Center for Policy Research Working Paper 15.*

Feng, Q., Kao, C., and Lazarova, S. (2009). Estimation of change points in panel models. Mimeo, Center for Policy Research, Syracuse University.

Han, A. K. and Park, D. (1989). Testing for structural change in panel data: Application to a study of u.s. foreign trade in manufacturing goods. *Review of Economics and Statistics*, 71:135–142.

Kiefer, J. (1959). K-Sample Analogues of the Kolmogorov-Smirnov and Cramer-V. Mises tests. *The Annals of Mathematical Statistics*, 30:420–447.

Lee, L. (2004). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica*, 72:1899–1925.

Lee, L. and Yu, J. (2010). Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, 154:165–185.

Liao, W. (2008). *Structural Breaks in Panel Data Models: A New Approach*. PhD thesis, New York University.

Pauwels, L. L., Chan, F., and Griffoli, T. M. (2012). Testing for structural change in heterogeneous panels with an application to the euro's trade effect. *Journal of Time Series Econometrics*.

Perron, P. (2006). Dealing with structural breaks. In Patterson, K. and Mills, T. C., editors, *Palgrave Handbook of Econometrics*, volume 1, pages 278–352. Palgrave Macmillan.

Qu, Z. and Perron, P. (2007). Estimating and testing structural changes in multivariate regressions. *Econometrica*, 75:459–502.

Wachter, S. and Tzavalis, E. (2012). Detection of structural breaks in linear dynamic panel data models. *Computational Statistics and Data Analysis*, 56:3020–3034.

Wied, D. (2013). Cusum-type testing for changing parameters in a spatial autoregressive model for stock returns. *Journal of Time Series Analysis*, 34:221–229.

Yu, J., Jong, R., and Lee, L. (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. *Journal of Econometrics*, 146.1:118–134.