Testing for Serial Correlation in Spatial Panels

Giovanni Millo

February 25, 2015

Abstract

We consider the issue of testing error persistence in spatial panels with individual heterogeneity. For random effects models, we review conditional Lagrange Multipliers tests from restricted models, and Likelihood Ratios or Wald tests via estimation of comprehensive models with correlation in space and time. We propose two ad-hoc tests for testing serial correlation in fixed effects panels, based either on time-demeaning or on forward orthogonal deviations. The proposed tests can be used under the RE assumption as well and are computationally less complicated than their RE counterparts. Both prove reasonably effective in our Montecarlo simulations.

1 Introduction

Panel data econometrics has been recently described as a "misspecification-test-free zone" (Banerjee *et al.*, 2010). This is generally not true for spatial panels, as the spatial econometrics literature has taken the utmost care in testing for spatial correlation. On the contrary, the applied literature on spatial panels has largely ignored serial correlation, as it has devoted limited attention to dynamic models: theoretical advances in either field have spanned few applications up to date. As Lee and Yu observe, "[i]n empirical applications with spatial panel data, it seems that investigators tend to limit their focus on some spatial structures and ignore others, and in addition, no serial correlation is considered" (Lee & Yu, 2012, p. 1370).

Methodologists have nevertheless considered estimation procedures allowing for serial error correlation in panel regression models with spatially autocorrelated outcomes and/or disturbances and random or fixed individual effects.

Baltagi *et al.* (2007) have extended the spatial panel framework to serial correlation in the remainder errors, while Elhorst (2008) has considered simultaneous error dependence in space and time. Lee & Yu (2012) proposed a very general specification including spatial lags, spatially and serially correlated errors together with individual effects. They assessed the biases due to neglecting serial correlation or some part of the spatial structure through Montecarlo simulation, and recommended a general to specific strategy.

Whether one approaches the issue of time persistence in spatial panels under the form of serial error correlation, or rather based on the specification of a dynamic model, testing for serial error correlation as a diagnostic check is

nevertheless relevant in both approaches, as any omitted dynamic would show up in error persistence. Moreover, very strong serial correlation, either in the form of an estimated parameter near 1 or of extremely high test statistics, can signal a nonstationarity problem, suggesting to reconsider the specification in a broader sense.

The issue of serial error correlation is particularly sensitive in the case of fixed effects panels, which on grounds of robustness are often the preferred alternative in many applied fields, as macroeconomics, regional applications or political science.

In fact, the standard technique for eliminating individual fixed effects, i.e. time-demeaning the variables, induces artificial serial correlation in the transformed residuals which can combine with the original correlation, if already present. By contrast, for pooled or random effects panels all three classic likelihood-based procedures are availeble. In a Lagrange Multipliers (LM) framework, one can use the C.2 test of BSJK (R implementation in package splm). The marginal/conditional version of the test assuming spatial but no random effects can be used to test in pooled models (although the C.2 version is still consistent if there are no random effects, so to stay on the safe side one can still use it). The comprehensive estimation framework for static panels described in Millo (2014) allows estimating both the general, encompassing model with both spatial and serial correlation, hence for likelihood ratio (LR) tests of the restriction of no serial correlation while allowing for spatial and/or random effects, i.e., for serial correlation testing of either RE or pooled models. Analogously, from within the encompassing model the significance diagnostics for the autoregressive parameter are equivalent to a Wald test for serial correlation.

The main contribution of the paper regards the proposal of a feasible strategy for testing serial error correlation in the case of fixed effects models, which is complicated by the ‘’artificial’ serial correlation induced by time-demeaning. In fact, if the original errors are serially uncorrelated, the transformed ones are negatively serially correlated with coefficient -1/(T-1). A Wooldridge-type test of serial correlation can then be based on an estimate of the serial correlation coefficient of the transformed model errors ψ:

- if the model is estimated by pooled or RE, testing the restriction $\psi = 0$
- if the model is estimated by FE, testing $\psi = -\frac{1}{T-1}$

Another possibility is to apply the alternative orthonormal transformation of Lee and Yu; the transformed residuals should then remain "white", so that the latter case reduces to the former.

In the first section of the paper we will set out the general model with unobserved heterogeneity and both spatially and serially correlated errors. Then we will review ML estimation of the model under the random effects hypothesis and describe the transformation approach to estimation in case the unobserved effects should be treated as fixed. A review of the well-known available tests of serial correlation for pooled or random effects spatial panels will then be followed by an outline of the novel strategy we propose for testing serial correlation under the fixed effects hypothesis. A Montecarlo exercise assessing the properties of our two proposed tests in small samples will follow.

2 Spatial panels with serial correlation

Following Baltagi *et al.* (2007), our point of departure is the following panel data regression model[1]

$$y_{it} = X_{it}'\beta + u_{it}, \qquad i = 1, \ldots, N, t = 1, \ldots, T \tag{1}$$

where y_{it} is the observation on cross-sectional unit i in time period t, and X_{it} is a $k \times 1$ vector of observations on the non-stochastic exogenous regressors. The disturbance vector is the sum of random regional effects and spatially autocorrelated residuals. In vector form this can be written as

$$u_t = \mu + \varepsilon_t \qquad \text{and} \qquad \varepsilon_t = \rho W \varepsilon_t + \nu_t. \tag{2}$$

The remaining disturbance term follows a first-order serially autocorrelated process

$$\nu_t = \psi \nu_{t-1} + e_t. \tag{3}$$

u_t, ε_t, ν_t and e_t are all $N \times 1$ columns vectors, μ is the random vector of $i.i.N(0, \sigma_\mu^2)$ region specific effects; ρ ($|\rho| < 1$) is the spatial autoregressive coefficient and ψ ($|\psi| < 1$) is the serial autocorrelation coefficient. As usual, W indicates the $N \times N$ matrix of known spatial weights whose diagonal elements are set to zero. $I_N - \rho W$ is assumed non-singular. Finally, $e_{it} \sim N(0, \sigma_e^2)$, $v_{i0} \sim N(0, \sigma_e^2/(1-\psi^2))$ and μ and ε are assumed to be independent.

The disturbance term can also be rewritten, in matrix notation, as

$$u = (\iota_T \otimes I_N)\mu + (I_T \otimes B^{-1})\nu \tag{4}$$

where $B = I_N - \rho W$, ι_T is a vector of ones, and I_T an identity matrix where T indicates the dimension. The model allows for serial correlation on each spatial unit over time as well as spatial dependence between spatial units at each time period. The presence of random effects accounts for possible heterogeneity across spatial units.

(Table 1 about here)

Depending on the restrictions on the parameters one can differently combine error features giving rise to various nested specifications (see Table 1). In particular, when both ψ and ρ are zero but σ_μ^2 is positive, the model reduces to a classical random effects panel data specification. When ρ is zero, the resulting model accounts for random effects with serially autocorrelated residuals. On the other hand, when ψ is zero and ρ and σ_μ^2 are not, the specification reduces to a random effects model with spatially autocorrelated residuals. Table 1 summarizes all possible specifications.

2.1 Estimation

We turn now to reviewing how to estimate the spatial panel model with serial correlation and either random or fixed effects, the results of which will be the basis for testing serial correlation.

Random effects models with spatial (SAR and/or SEM) correlation and serial correlation in the remainder error can be estimated by maximum likelihood.

[1]The extension to a spatially lagged dependent variable is described in Millo (2014), to which the reader is referred.

If the exogeneity assumption for individual effects does not hold ('fixed effects' case), then the latter have to be eliminated before estimation. This is usually accomplished by either differencing or time-averaging. One further transformation is discussed below. Once the fixed effects have been transformed out, a restricted version of the above model assuming $\mu_i = 0 \forall i$ can be estimated.

ML estimation of an encompassing model In the present section we first discuss the estimation approach to the richest specification, i.e. the one allowing for random effects, serial and spatial correlation. The special case without random effects will be discussed subsequently.

To derive the expression for the likelihood, Baltagi *et al.* (2007) use a Prais-Winsten transformation of the model with random effects and spatial autocorrelation. Following their simplifying notation, define

$$
\begin{aligned}
&\alpha = \sqrt{\tfrac{1+\psi}{1-\psi}} \\
&d^2 = \alpha^2 + (T-1) \\
&V_\psi = \tfrac{1}{1-\psi^2} V_1 \\
&V_1 = \begin{bmatrix} 1 & \psi & \psi^2 & \dots & \psi^{T-1} \\ \psi & 1 & \psi & \dots & \psi^{T-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi^{T-1} & \psi^{T-2} & \psi^{T-3} & \dots & 1 \end{bmatrix}.
\end{aligned}
$$

Scaling the error covariance matrix by the idiosyncratic error variance σ_ε^2 and denoting $\phi = \frac{\sigma_\mu^2}{\sigma_\varepsilon^2}$, the expressions for the scaled error covariance matrix Σ and for its inverse Σ^{-1} and determinant $|\Sigma|$ can be written respectively as

$$
\begin{aligned}
&\Sigma = \phi(J_T \otimes I_N) + V_\psi \otimes (B'B)^{-1} \\
&\Sigma^{-1} = V_\psi^{-1} \otimes (B'B) + \tfrac{1}{d^2(1-\psi)^2}(V_\psi^{-1} J_T V_\psi^{-1}) \\
&\qquad \otimes ([d^2(1-\psi)^2 \phi I_N + (B'B)^{-1}]^{-1} - B'B) \\
&|\Sigma| = |\Sigma^*|/(1-\psi^2)^N.
\end{aligned}
$$

with $|\Sigma^*| = |d^2(1-\psi)^2 \phi I_N + (B'B)^{-1}| \cdot |(B'B)^{-1}|^{T-1}$. Therefore, one can derive the expression of the likelihood:

$$
\begin{aligned}
L(\beta, \sigma_e^2, \phi, \psi, \rho) = &-\tfrac{NT}{2} 2\pi - \tfrac{NT}{2} \ln \sigma_e^2 + \tfrac{N}{2} \ln(1-\psi^2) \\
&- \tfrac{1}{2} \ln |d^2(1-\psi)^2 \phi I_N + (B'B)^{-1}| \\
&+ (T-1) \ln |B| - \tfrac{1}{2\sigma_e^2} u' \Sigma^{-1} u
\end{aligned}
$$

Millo (2014) describes the iterative procedure to obtain the maximum likelihood estimates in the extended model comprising a spatial lag of the dependent variable. Starting from initial values for ρ, ψ and ϕ, one can obtain estimates for β and σ_e^2 from the first order conditions:

$$\beta = (X'\Sigma^{-1}X)^{-1} X'\Sigma^{-1} y \tag{5}$$

$$\sigma_e^2 = (y - X\beta)'\Sigma^{-1}(y - X\beta)/NT. \tag{6}$$

The likelihood can be concentrated and maximized with respect to ρ, ψ and ϕ. The estimated values of ρ, ψ and ϕ are then used to update the expression for Σ^{-1}. These steps are then repeated until convergence. In other words, for a

specific Σ the estimation can be operationalized by a two steps iterative procedure that alternates between GLS (for β and σ_e^2) and concentrated likelihood (for the remaining parameters) until convergence.

Statistical inference can then be based on the expression of the information matrix. Millo & Piras (2012); Millo (2014) obtain standard errors for β from GLS, and employ a numerical Hessian to perform statistical inference on the error components.

These steps remain valid when the model to be estimated is one of the reduced forms presented in Table 1. In particular, the specification where $\phi = 0$ will be of interest in our case.

The estimate of the comprehensive model can be the basis for either a direct assessment of the magnitude and significance of the serial correlation coefficient, or more modestly (as will be the case for fixed effects procedures) for a serial correlation test.

Estimating the FE model by transformation In this section we review the general transformation approach to the estimation of panel models with fixed effects; then its application to spatial panel models.

If the individual effect cannot be assumed independent from the regressors, fixed effects (FE) methods are in order. The modern approach to the issue, tracing back to Mundlak (1978) and summarized, among others, in Wooldridge (2002, 10.2.1), centers on the statistical properties of the individual effects. If uncorrelated, then individual effects can be considered as a component of the error term, and treated in a generalized least squares fashion as seen above. If not, then the latter strategy leads to inconsistency; the individual effects will have to be estimated or, more frequently, eliminated by first differencing or time-demeaning the data (see Wooldridge, 2002, 10.5).[2]. In a spatial setting, Lee & Yu (2012) give an extensive treatment to which the reader is referred here.

The well-known time-demeaning, or *within* transformation, entails subtracting averages over the time dimension, so that the model becomes:

$$y_{it} - \bar{y}_i = (X_{it} - \bar{X}_i)\beta + (u_{it} - \bar{u}_i) \tag{7}$$

where $\bar{y}$ and $\bar{X}$ denote time means of y and X

From a computational viewpoint, according to the framework of Elhorst (2003), fixed effects estimation of spatial panel models is accomplished as pooled estimation on time-demeaned data. Hence, it is fully encompassed by the method outlined in the previous section, but for the fact that individual effects can now be omitted.

Elhorst's procedure has long been the standard in applied practice and available software, but has been questioned by Anselin *et al.* (2008) because time-demeaning alters the properties of the joint distribution of errors, introducing serial dependence: see Lee & Yu (2010b, p.257) for a discussion of the issue, and Millo & Piras (2012, p.33) for an evaluation if its practical significance through Montecarlo simulation. As it turns out, the transformation induces bias only in the estimate of the errors' variance, while those of the regressors' coefficients β and the spatial coefficients (λ, ρ) remain consistent. Hence, the residuals are still pointwise consistent estimates of the errors.

[2]A short introduction with the basic references can be found in Baltagi (2008b, 2.3.1)

Nevertheless, their dispersion is biased while ideally one would want to perform inference on the basis of an unbiased estimate of both error mean and variance. To solve the problem, Lee & Yu (2010a, 3.2) suggest either a correction ex-post or to apply to spatial data a different transformation:

$$Oy_{it} = OX_{it}\beta + Ou_{it} \tag{8}$$

where

$$O = \begin{pmatrix} \sqrt{\frac{T-1}{T}} & -\frac{1}{\sqrt{T(T-1)}} & -\frac{1}{\sqrt{T(T-1)}} & \cdots & -\frac{1}{\sqrt{T(T-1)}} & -\frac{1}{\sqrt{T(T-1)}} \\ 0 & \sqrt{\frac{T-2}{T-1}} & -\frac{1}{\sqrt{(T-1)(T-2)}} & \cdots & -\frac{1}{\sqrt{(T-1)(T-2)}} & -\frac{1}{\sqrt{(T-1)(T-2)}} \\ 0 & 0 & \sqrt{\frac{T-3}{T-2}} & \cdots & -\frac{1}{\sqrt{(T-2)(T-3)}} & -\frac{1}{\sqrt{(T-2)(T-3)}} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \sqrt{\frac{1}{2}} & -\sqrt{\frac{1}{2}} \end{pmatrix}$$

This orthonormal transformation is well known from the panel data literature as *forward orthogonal deviations* (henceforth OD) (see Arellano, 2003, p. 17) and is justified as the GLS transformation to remove MA(1) correlation from first-differenced data (ibid.). From our viewpoint, it has the desirable characteristic of not inducing any serial correlation in the transformed errors. Moreover, as far as β is concerned, the estimator resulting from applying OLS to the OD transformed data gives the same result as the FE one, so that $\hat{\beta}_{OD} = \hat{\beta}_{FE}$.

The OD transformation can be employed in the estimation of spatial panel fixed effects models as an alternative procedure w.r.t. demeaning and then correcting the variance (see Lee & Yu, 2010a, 3.2). The application to models with serial correlation is still undocumented and is left for future work; here we will employ a combination of the serial-spatial estimator outlined above and the OD transformation as a testing device, based on the fact that if the original errors are serially incorrelated, then the OD transformed ones must still be.

3 Testing for serial correlation in spatial RE panels

In this section we review existing testing procedures for serial correlation in spatial (SAR and/or SEM) panels with uncorrelated heterogeneity. As it turns out, all three likelihood-based standard procedures are available: Wald, likelihood ratio (LR) and Lagrange multiplier (LM) tests; the latter, though, has not been derived for models containing a spatial lag (SAR).

Spatial panels without individual effects, or with individual idiosyncracies that comply with the random effects hypothesis, can be estimated in the most comprehensive of the above specifications and then from here the zero restriction on the autoregressive coefficient can be tested, as in a Wald-type procedure. Alternatively, an asymptotically equivalent procedure can be employed: estimating the reduced model as well and then Likelihood Ratio (LR) testing one versus the other. As a last option, the third likelihood-based procedure can be employed: the Lagrange multiplier (LM) testing procedure, also known as *score test*, which is based on verifying whether the score of the likelihood of a

restricted model is significantly different from the zero vector. If not, then the restriction is not binding w.r.t. the problem at hand and it is thus accepted . With respect to its asymptotically equivalent siblings, the Likelihood Ratio and Wald tests, the LM test requires only estimation of the restricted model. In this section we review the joint J and conditional $C.3$ tests for serial correlation by Baltagi *et al.* (2007). The hypotheses under consideration are:

1. $H_0^a : \rho = \psi = \sigma_\mu^2 = 0$ under under the alternative that at least one component is not zero (J)
2. $H_0^i : \psi = 0$, assuming $\rho \neq 0, \sigma_\mu^2 > 0$: test for serial correlation, allowing for spatial correlation and random individual effects ($C.2$)

The joint LM test for testing H_0^a is given by:

$$LM_j = \frac{NT^2}{2(T-1)(T-2)}[A^2 - 4AF + 2TF^2] + \frac{N^2T}{b}H^2 \tag{9}$$

where, $A = \tilde{u}'(J_T \otimes I_N)\tilde{u}/\tilde{u}'\tilde{u} - 1$, $F = \tilde{u}'(G_T \otimes I_N)\tilde{u}/2\tilde{u}'\tilde{u}$, $H = \tilde{u}'(I_T \otimes (W' + W))\tilde{u}/2\tilde{u}'\tilde{u}$, $b = tr(W + W')^2/2$, G is a matrix with bidiagonal elements equal to one and $\tilde{u}$ denotes OLS residuals. Under H_0^a, LM_J is distributed as χ_3^2.

The conditional $C.2$ test for H_o^i is based on the following statistic, asymptotically distributed as χ_1^2 under H_0^i:

$$LM_{\psi/\rho\mu} = \hat{D}(\psi)^2 J_{33}^{-1} \tag{10}$$

where J_{33}^{-1} is the $(3,3)$ element of the information matrix given in Baltagi *et al.* (2007) (equation 3.10);

$$\begin{aligned}\hat{D}(\psi) = & -\frac{T-1}{T}(\hat{\sigma}_e^2 tr(Z(B'B)^{-1}) - N) \\ & + \frac{\hat{\sigma}_e^2}{2}\hat{u}'[\frac{1}{\sigma_e^4}(E_T G E_T) \otimes (B'B) + \frac{1}{\sigma_e^2}(\bar{J}_T G E_T) \otimes Z \\ & + \frac{1}{\sigma_e^2}(E_T G \bar{J}_T) \otimes Z + (\bar{J}_T G \bar{J}_T) \otimes Z(B'B)^{-1}Z]\hat{u}\end{aligned} \tag{11}$$

with $Z = [T\sigma_\mu^2 I_N + \sigma_e^2(B'B)^{-1}]$ is the score under the null hypothesis and $\hat{u}$ the vector of residuals under the null, i.e., from ML estimation of the panel model with individual error components and serial correlation in idiosyncratic errors;

$$g = \frac{1}{\sigma_e^2}(1-\psi)2 + (T-2)(1-\psi) \tag{12}$$

and b has been defined above. It shall be observed that while the J test only needs OLS residuals and is therefore computationally very simple, but does not give any information about which of the three possible effects is actually present, the more interesting $C.3$ test needs residuals from a specification which is much more difficult to compute.

4 Testing for serial correlation in the presence of fixed effects

Testing for serial correlation in spatial panels with fixed effects is undocumented, to the best of our knowledge. Yet all devices needed are already established in the econometrics literature and ready to be combined. The procedures we propose are based on the general approach of Wooldridge (2002), who suggests to run an autoregression on the residuals of a fixed effects model and check whether the relevant coefficient is statistically different from its expected value under the null hypothesis of no serial correlation. The latter is, depending on the transformation used for eliminating the fixed effects, either a function of the time dimension T in the case of time-demeaning, or zero if using forward orthogonal deviations as described below. We propose the application of a similar principle to spatial panels.

4.1 General serial correlation tests

A general testing procedure for serial correlation in fixed effects (FE), random effects (RE) and pooled-OLS panel models alike can be based on considerations in (Wooldridge, 2002, 10.7.2). For the random effects model, he observes that under the null of homoskedasticity and no serial correlation in the idiosyncratic errors, the residuals from the quasi-demeaned regression must be spherical as well. Else, as the individual effects are wiped out in the demeaning, any remaining serial correlation must be due to the idiosyncratic component. Hence, a simple way of testing for serial correlation is to apply a standard serial correlation test to the quasi-demeaned model. The same applies in a pooled model, w.r.t. the original data.

The FE case is different. It is well-known that if the original model's errors are uncorrelated then FE residuals are negatively serially correlated, with $cor(\hat{u}_{it}, \hat{u}_{i,t-1}) = -1/(T-1)$ for each t (see Wooldridge, 2002, 10.5.4). This correlation disappears as T diverges, so this kind of test is readily applicable to time-demeaned data only for T "sufficiently large". Baltagi and Li derive a basically analogous T-asymptotic test for first-order serial correlation in a FE panel model as a Breusch-Godfrey LM test on within residuals (see Baltagi & Li, 1995, par. 2.3 and formula 12). They also observe that the test on within residuals can be used for testing on the RE model, as "the within transformation [time-demeaning, in our terminology] wipes out the individual effects, whether fixed or random", a consideration we will recall in the following. On a related note, generalizing the Durbin-Watson test to FE models by applying it to fixed effects residuals is documented in Bhargava *et al.* (1982).

For the reasons reported above, under the null of no serial correlation in the errors, the residuals of a FE model must be negatively serially correlated, with coefficient equal to $-1/(T-1)$. Wooldridge suggests basing a test for this null hypothesis on a pooled regression of FE residuals on themselves, lagged one period:

$$\hat{\epsilon}_{i,t} = \alpha + \delta\hat{\epsilon}_{i,t-1} + \eta_{i,t}$$

Rejecting the restriction $\delta = -1/(T-1)$ makes us conclude against the original null of no serial correlation. A Wooldridge-type test of serial correlation can then

be based on an estimate of the serial correlation coefficient of the transformed model errors ψ:

- if the model is estimated by pooled or RE, testing the restriction $\psi = 0$
- if the model is estimated by FE, testing $\psi = -\frac{1}{T-1}$

Another possibility is to apply the alternative orthonormal transformation of Lee and Yu; the transformed residuals should then remain "white", so that the latter case reduces to the former.

4.2 Two serial correlation tests for spatial FE panels

By analogy, the problem of testing for serial correlation in the residuals of *spatial* panels can be addressed combining the above testing framework with the comprehensive estimation approach including serial correlation described in precedence; but obviously omitting the random effects features, so that the likelihood simplifies to:

$$\begin{aligned} L(\beta, \sigma_e^2, \psi, \rho) = -\tfrac{NT}{2} 2\pi - \tfrac{NT}{2} \ln \sigma_e^2 + \tfrac{N}{2} \ln(1 - \psi^2) \\ + T \ln |B| - \tfrac{1}{2\sigma_e^2} u' \Sigma^{-1} u \end{aligned}$$

and

$$\Sigma^{-1} = V_\psi^{-1} \otimes (B'B)$$

considerably simplifying the numerical estimation procedure, especially as one does not need to calculate $(B'B)^{-1}$ any more, but also because V_ψ^{-1} has a convenient self-similar representation (see Millo, 2014, 5.1). This model can be estimated on relatively big samples and the optimization of its likelihood turns out computationally simpler than that of the spatial random effects model whose residuals are needed as the basis for the conditional C.2 test of Baltagi *et al.* (2007) (see Millo, 2014, Table 2). Hence employing a FE-type test based on elimination of the individual heterogeneity, although suboptimal under RE, can turn out to be both safer than the RE-type procedures as the underlying hypotheses are concerned, and computationally less burdensome.

Two different tests can be performed on the estimates, depending on the way individual effects have been transformed out.

Wooldridge-type AR test A Wooldridge-type AR(1) test for spatial panels of either SAR, SEM or SAREM type can be based on testing the derived null hypothesis $H_0 : \psi = -\frac{1}{T-1}$ in the full model estimated on time demeaned data.

The test, which we here label AR_{FE}, will be appropriate for any T, and particularly for short panels.

Orthogonal-deviation based AR test An equivalent test, henceforth AR_{OD}, can be based on testing the more familiar hypothesis $H_0 : \psi = 0$ if the data are transformed through the non-correlation-inducing forward orthogonal deviations transformation. As T diverges, the induced correlation in the time-demeaning case tends to zero and the difference between the two procedures wanes.

In the following we will ask which one performs better in real-world conditions and provide a first answer through Montecarlo simulation. But first let us go through a short illustration through a well-known example.

4.3 Illustration

To illustrate the use (and the relevance) of the different tests we will resort to a well-known dataset which has been recently employed in a number of spatial econometric studies (Baltagi & Li, 2004; Elhorst, 2005, 2012; Kelejian & Piras, 2011; Debarsy *et al.*, 2012; Vega & Elhorst, 2013; Kelejian & Piras, 2014) and will therefore be familiar to most researchers.

The Cigarette dataset is taken from Baltagi (2008a)[3]; the original application is in Baltagi & Levin (1992). Further reconsiderations include Baltagi *et al.* (2000); Baltagi & Griffin (2001). It contains data for the years 1963-1992 and 46 American states on: real per capita sales of cigarettes per person of smoking age (i.e., over 14) measured in packs (C), average real retail price per pack (P), real disposable income per capita (Y) and the minimum price per pack in neighbouring states (Pn). The last variable is included in the original application in order to proxy for cross-border smuggling (Baltagi, 2008a, p. 156); this could also be done controlling for spatial effects, as in the above mentioned cases.

Individual (state-specific) effects are included to account for idiosyncratic characteristics of territory, like the presence of tax-exempt military bases or indian reservations, the prevalence of a religion that forbids smoking (the Mormons in Utah) or the effect of tourism. Time effects are also included to account for (USA-wide) policy interventions and warning campaigns. Given their peculiar nature, both kinds of effects will better be assumed fixed; but in the following we consider both testing under the RE and under the FE hypothesis.

The original application is dynamic, as it contains lagged consumption in order to control for habit persistence in smoking. Nevertheless, a static version of the Cigarette model has often been employed:

$$lnC_{it} = \alpha + \beta_1 lnP_{it} + \beta_2 lnY_{it} + u_{it}$$

and this latter we will use in our case; given the theoretical reasons for persistence, it will be of particular importance to test for serial error correlation.

In the following Table 7, all tests mentioned in the paper are performed on the static spatial error specification:

(Table 6 about here)

Testing the comprehensive SAR+SEM model gives similar results (omitted), but for the fact that the LM_j and $LM_{C.2}$ tests cannot be employed any more.

If the random effects hypothesis can be trusted, the left part of the table can be considered: LM, LR and Wald-type tests and the estimate of ψ. All likelihood-based tests signal a very strong departure from the null of serial incorrelation, while $\hat{\psi}$ from estimation of the comprehensive model is very near to one. The AR_{OD} and AR_{FE} tests also reject the null well beyond any conventional significance level.

If the random effects hypothesis is considered dubious, then one can only look at the results from the AR_{OD} and AR_{FE} tests; again, there is little doubt about high persistence in the error terms.

In any case, the results point to very strong autoregressive behaviour in the residuals, bordering with unit roots if we believe the RE hypothesis: a static spatial panel specification assuming timewise-incorrelated errors is inappropriate, and an analysis of stationarity would be advisable.

[3]The spatial weights matrix is due to Paul Elhorst; data and weights can be found, respectively, in the R packages Ecdat (Croissant, 2010) and splm (Millo & Piras, 2012).

Lastly, computing times for the AR_{OD} and AR_{FE} tests are smaller than those for their LM, LR and Wald counterparts by an order of magnitude; although all of them are still feasible on any machine for this - rather moderate - sample size.

5 Montecarlo experiments

The properties of serial correlation tests in spatial RE models have already been established by extensive simulations in Baltagi *et al.* (2007). Here we consider only our new two procedures for testing under FE. In the Montecarlo experiments, we consider the rejection rates of either test at the 5 percent significance level. This gives an assessment of the empirical size if the data are simulated under the null hypothesis of no serial correlation, and of the empirical power of the test under alternative data generating processes where $\psi \neq 0$.

The simulated idiosyncratic innovations are distributed as a standard Normal, and the individual effects as $N(0, \mu)$, so that μ is the ratio of error variances. Along with an intercept term, we consider two regressors: x_1 is sampled from a Uniform [-7.5, 7.5], x_2 is drawn from a standard Normal. The simulation parameters are chosen with a target R^2 of 0.7. The coefficients for the regressors are set to 0.5 and 10, respectively. Our spatial layout is given by the 48 states of the continental US. The spatial weighting matrix is a simple binary contiguity one. We consider two values for the number of time periods, one representative of a typical "short" panel, the other of macroeconomic panels found in the literature, and set $T = 4, 15$. We allow combinations of two different values for both ρ and ψ, namely either zero (no effect) and 0.5, so that next to the usual two cases of spatial lag (SAR) and spatial error (SEM) we consider both the case of no spatial correlation and that of combined SAR and SEM processes. We consider three values for the objective parameter ψ: zero, corresponding to no error persistence, and two positive levels of serial correlation: 0.3 (weak) and 0.8 (strong). For all experiments 1,000 replications are performed.

Simulation results are reported below in Tables 2 to 4 .

[Tables 2,3,4 about here]

Test size under validity of the null is reasonably close to 5% in all experiments, more so when $T = 15$ than in the short panel. Empirical power is very good for the long panel both for weak ($\psi = 0.3$) and for strong ($\psi = 0.8$) error persistence; in the short panel case, power is moderate (near 50 %) for the OD-based test, while lower for the Wooldridge-type variant. Results are much the same under either type of spatial dependence, both or none alike; and under presence or absence of individual effects, testifying how effectively the procedure controls for spatial features. The AR_{OD} test uniformly dominates the Wooldridge-type version AR_{FE} over the short sample; for the longer panel, the results of the former are still slightly better, but in this case performance is satisfactory for both versions so that either can be safely employed.

6 Conclusions

We address the much neglected issue of testing for serial error correlation in spatial panels of lag or error type or both, possibly containing individual heterogeneity of the random or fixed effects type.

Comprehensive estimators both for the encompassing model and for its restrctions have been developed for the random effects case, as well as joint and conditional Lagrange multiplier tests, so that the zero-restriction of the serial correlation coefficient can be tested by either of the three well-known likelihood-based procedures: Wald test (corresponding to the significance diagnostics of the serial correlation coefficient in the full model), likelihood ratio test based on the difference between the full and restricted models' log-likelihoods, and Lagrange multiplier test based on the restricted model only, as derived by Baltagi *et al.* (2007). We have briefly reviewed them here.

In contrast to this wealth of techniques, for which user-friendly software is freely available to researchers, the fixed effects case – which is both considered the most interesting one in spatial applications (see the discussion in the Introduction) and is also robust (although less efficient) in case the random effect assumption should hold true – has neither seen practical applications, nor methodological attention.

We propose two feasible procedures for the fixed effects case, one based on observations in Wooldridge (2002) and the other on the work of Lee & Yu (2010a). The former consists in estimating the full spatial model on time-demeaned data and testing the resulting serial correlation coefficient for departures from the implied negative serial correlation induced by the demeaning transformation; the latter in employing the forward orthogonal deviations transformation of Arellano (2003) instead of time-demeaning, which maintains the original correlation properties in transformed residuals, and directly testing the resulting coefficient for departures from zero.

Elimination of the individual effects through transformation is appropriate, although statistically suboptimal, even if the RE hypothesis holds; hence our proposed FE-type tests, unlike the RE ones, can be safely employed in dubious situations. Moreover, the likelihood optimization procedure they are based upon is considerably simpler than that of the full model with RE, and has been proven to work even on relatively big samples (see Millo, 2014, Table 2). The computational burden from performing the proposed tests is actually smaller than that of the conditional C.2 test of Baltagi *et al.* (2007), which requires to estimate a spatial model with random effects.

A short Montecarlo experiment illustrates the size and power properties of the two proposed procedures, which turn out satisfactory for both tests when $T \geq 10$ while the OD-based test fares better than the FE-based one in the short panel case ($T = 4$).

7 Appendix: Computational details

All the procedures in the paper are available through user-friendly R implementations, some of which are forthcoming and can be requested to the contact author. In particular, Baltagi *et al.* (2007) tests are available as different options of the function `bsjktest` in package **splm** (Millo & Piras, 2012). LR

and Wald tests depend on the estimation of restricted and unrestricted models through the function `spreml` as extensively documented in Millo (2014); the same goes for numerical estimation of the serial correlation coefficient ψ. Lastly, the AR_{FE} and AR_{OD} tests can be performed by combining the `Within` and `Orthog` transformation functions from the general-purpose panel data package **plm** (Croissant & Millo, 2008) - the second of which is forthcoming and can be requested to the contact author - and `spreml` described above.

Functionality is summarized in Table 7.

Table 1: Different model specifications that can be generated as special cases of the general specification.

	$\rho \neq 0$ $\rho \neq 0$	$\rho \neq 0$ $\rho = 0$	$\rho = 0$ $\rho \neq 0$	$\rho = 0$ $\rho = 0$
$\sigma^2_\mu \neq 0$	SEMSRRE	SEMRE	SSRRE	RE
$\sigma^2_\mu = 0$	SEMSR	SEM	SSR	OLS

Table 2: Empirical size of test for different spatial processes

AR_{OD}						
		σ_μ	0	0	1	1
		ρ	0	0.5	0	0.5
T	λ					
4	0		0.062	0.061	0.068	0.069
4	0.5		0.064	0.066	0.063	0.061
10	0		0.057	0.052	0.052	0.051
10	0.5		0.045	0.056	0.050	0.060
15	0		0.067	0.052	0.044	0.055
15	0.5		0.041	0.054	0.054	0.051
AR_{FE}						
		σ_μ	0	0	1	1
		ρ	0	0.5	0	0.5
T	λ					
4	0		0.03	0.024	0.068	0.069
4	0.5		0.04	0.032	0.063	0.061
10	0		0.049	0.057	0.052	0.051
10	0.5		0.037	0.044	0.050	0.060
15	0		0.045	0.045	0.044	0.055
15	0.5		0.038	0.048	0.054	0.051

Table 3: Empirical size and power

AR_{OD}

		σ_μ	0	0	0	0	1	1	1	1
		ρ	0	0	0.5	0.5	0	0	0.5	0.5
		λ	0	0.5	0	0.5	0	0.5	0	0.5
T	ψ									
4	0		0.062	0.061	0.064	0.066	0.068	0.069	0.063	0.061
4	0.3		0.459	0.453	0.429	0.490	0.442	0.471	0.472	0.460
4	0.8		0.993	0.990	0.996	0.994	0.995	0.991	0.993	0.996
10	0		0.057	0.052	0.045	0.056	0.052	0.051	0.050	0.060
10	0.3		0.998	0.999	0.997	0.998	0.999	0.999	0.998	0.998
10	0.8		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
15	0		0.067	0.052	0.041	0.054	0.044	0.055	0.054	0.051
15	0.3		0.997	1.000	0.998	0.999	0.999	0.999	0.998	1.000
15	0.8		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4: Empirical size and power

AR_{FE}

		σ_μ	0	0	0	0	1	1	1	1
		ρ	0	0	0.5	0.5	0	0	0.5	0.5
		λ	0	0.5	0	0.5	0	0.5	0	0.5
T	ψ									
4	0 4		0.030	0.024	0.040	0.032	0.020	0.034	0.034	0.042
4	0.3		0.276	0.274	0.253	0.272	0.257	0.262	0.271	0.261
4	0.8		0.964	0.956	0.962	0.957	0.965	0.962	0.966	0.961
10	0		0.049	0.057	0.037	0.044	0.039	0.049	0.049	0.038
10	0.3		1.000	0.999	0.998	0.999	1.000	1.000	0.998	0.998
10	0.8		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
15	0		0.045	0.045	0.038	0.048	0.037	0.042	0.046	0.051
15	0.3		0.997	1.000	1.000	0.999	0.999	0.998	0.999	1.000
15	0.8		1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 5: Serial correlation tests, SEM specification

	LM_j	$LM_{C.2}$	LR	$Wald$	$\hat{\psi}$	AR_{OD}	AR_{FE}
Statistic	12588.9	885.2	2034.7	286.0	0.98	86.6	89.7
Distribution	χ^2_3	χ^2_1	χ^2_1	z	-	z	z
p-value	0	0	0	0	-	0	0
Computing time	0.67	63.68	113.49	47.84	-	3.19	3.17

Table 6: Available user-level functionality

	LM_j	$LM_{C.2}$	LR	$Wald$	$\hat{\psi}$	AR_{OD}	AR_{FE}
Function	`bsjktest`	`bsjktest`	`lrtest`	`spreml`	`spreml`	forthcoming	forthcoming
Option	`test='J'`	`test='C.2'`	n.a.	n.a.	n.a.	-	-
In package	**splm**	**splm**	**splm**	**splm**	**splm**	-	-

References

Anselin, L., Le Gallo, J., & Jayet, H. 2008. Spatial Panel Econometrics. *Pages 624 – 660 of:* Matyas, L., & Sevestre, P. (eds), *The econometrics of Panel Data, Fundamentals and Recent Developments in Theory and Practice (3rd Edition)*. Springer-Verlag, Berlin Heidelberg.

Arellano, Manuel. 2003. Panel data econometrics. *OUP Catalogue*.

Baltagi, Badi. 2008a. *Econometric analysis of panel data*. Vol. 1. John Wiley & Sons.

Baltagi, Badi H, & Griffin, James M. 2001. The econometrics of rational addiction: the case of cigarettes. *Journal of Business & Economic Statistics*, **19**(4), 449–454.

Baltagi, Badi H, & Levin, Dan. 1992. Cigarette taxation: raising revenues and reducing consumption. *Structural Change and Economic Dynamics*, **3**(2), 321–335.

Baltagi, Badi H, & Li, Dong. 2004. Prediction in the panel data model with spatial correlation. *Pages 283–295 of: Advances in spatial econometrics*. Springer.

Baltagi, Badi H, & Li, Qi. 1995. Testing AR (1) against MA (1) disturbances in an error component model. *Journal of Econometrics*, **68**(1), 133–151.

Baltagi, Badi H, Griffin, James M, & Xiong, Weiwen. 2000. To pool or not to pool: Homogeneous versus heterogeneous estimators applied to cigarette demand. *Review of Economics and Statistics*, **82**(1), 117–126.

Baltagi, B.H. 2008b. *Econometric Analysis of Panel Data, 4th edition*. New York: Wiley.

Baltagi, B.H., Song, S.H., Jung, B.C., & Koh, W. 2007. Testing for serial correlation, spatial autocorrelation and random effects using panel data. *Journal of Econometrics*, **140**(1), 5–51.

Banerjee, Aninday, Eberhardt, Markus, & Reade, J James. 2010. *Panel Estimation for Worriers*. Tech. rept. Department of Economics, University of Birmingham.

Bhargava, Alok, Franzini, Luisa, & Narendranathan, Wiji. 1982. Serial correlation and the fixed effects model. *The Review of Economic Studies*, **49**(4), 533–549.

Croissant, Yves. 2010. *Ecdat: Data sets for econometrics*. R package version 0.1-6.

Croissant, Yves, & Millo, Giovanni. 2008. Panel data econometrics in R: The plm package. *Journal of Statistical Software*, **27**(2), 1–43.

Debarsy, Nicolas, Ertur, Cem, & LeSage, James P. 2012. Interpreting dynamic space–time panel data models. *Statistical Methodology*, **9**(1), 158–171.

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

Elhorst, J Paul. 2012. Matlab software for spatial panels. *International Regional Science Review*, 0160017612452429.

Elhorst, J.P. 2003. Specification and estimation of spatial panel data models. *International Regional Sciences Review*, **26**(3), 244–268.

Elhorst, J.P. 2008. Serial and Spatial error correlation. *Economics Letters*, **100**(3), 422–424.

Kelejian, Harry H, & Piras, Gianfranco. 2011. An extension of Kelejian's J-test for non-nested spatial models. *Regional Science and Urban Economics*, **41**(3), 281–292.

Kelejian, Harry H, & Piras, Gianfranco. 2014. Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. *Regional Science and Urban Economics*, **46**, 140–149.

Lee, L., & Yu, J. 2012. Spatial panels: random components versus fixed effects. *International Economic Review*, **53**(4), 1369–1412.

Lee, L.F., & Yu, J. 2010a. Estimation of spatial autoregressive panel data models with fixed effects. *Journal of Econometrics*, **154**(2), 165–185.

Lee, L.F., & Yu, J. 2010b. Some recent developments in spatial panel data models. *Regional Science and Urban Economics*, **40**(5), 255–271.

Millo, Giovanni. 2014. Maximum likelihood estimation of spatially and serially correlated panels with random effects. *Computational Statistics & Data Analysis*, **71**, 914–933.

Millo, Giovanni, & Piras, Gianfranco. 2012. splm: Spatial panel data models in R. *Journal of Statistical Software*, **47**(1), 1–38.

Mundlak, Y. 1978. On the pooling of time series and cross-section data. *Econometrica*, **46**, 69–85.

Vega, Solmaria Halleck, & Elhorst, J Paul. 2013. On spatial econometric models, spillover effects, and W. *In: 53rd ERSA conference, Palermo.*

Wooldridge, J.M. 2002. *Econometric analysis of cross-section and panel data.* MIT Press.