

Millo, Giovanni

Conference Paper

Testing Tobler's law in spatial panels: a test for spatial dependence robust against common factors

50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Millo, Giovanni (2010) : Testing Tobler's law in spatial panels: a test for spatial dependence robust against common factors, 50th Congress of the European Regional Science Association: "Sustainable Regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/119050>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Testing Tobler's law in spatial panels: a test for spatial dependence robust against common factors

Giovanni Millo

June 19, 2010

Abstract

In the spatial econometrics literature, spatial error dependence is characterized by spatial autoregressive processes, which relate every observation in the cross-section to any other with distance-decaying intensity: i.e., dependence obeys Tobler's First Law of Geography ("everything is related to everything else, but near things are more related than distant things"). In the literature on factor models, on the converse, the degree of correlation between cross-sectional units depends only on factor loadings. Standard spatial correlation tests have power against both types of dependence, while the economic meaning of the two can be much different; so it may be useful to devise a test for detecting "distance-related" dependence in the presence of a "factor-type" one. Pesaran's CD is a test for global cross-sectional dependence with good properties. The CD(p) variant only takes into account p-th order neighbouring units to test for local cross-sectional dependence. The pattern of CD(p) as p increases can be informative about the type of dependence in the errors, but the test power changes as new pairs of observations are taken into account. I propose a bootstrap test based on the values taken by the CD(p) test under permutations of the neighbourhood matrix, i.e. when "resampling the neighbours". I provide Montecarlo evidence of it being able to tell the presence of spatial-type dependence in the errors of a typical spatial panel irrespective of the presence of an unobserved factor structure.

1 Introduction

The concepts of local and global spatial dependence are cornerstones of regional economics and spatial econometrics, while an ever growing strand of the econometric literature is dedicated to the issues of estimation and testing under cross-sectional dependence at large.

In the literature on spatial econometrics, *local* dependence is usually characterized by moving average spatial processes, which relate every observation to its neighbours only; *global* dependence by spatial autoregressive processes, which can be shown to relate every observation in the cross-section to any other, although with intensity decaying with the distance between them. In both cases, dependence obeys Tobler's First Law of Geography, stating that

everything is related to everything else, but near things are more related than distant things (Tobler 1970).

In the literature on factor models, on the converse, the degree of correlation between cross-sectional units depends on factor loadings, so distance doesn't play any role¹. Far away observations can react to a common shock in the same way as nearby ones or not, depending on the individual loadings.

While spatial econometric models allow estimation on single cross-sections, factor models are to be estimated on panel data. For these, estimators that are consistent in presence of either type of dependence have been devised, so that for the sake of estimation discriminating between the two may in some cases not be much of a problem. Nevertheless, consistent estimators for general cross-sectional dependence like Pesaran's CCEMG ([3]) still rely on a "biggish" time dimension, while the ability to characterize dependence in a parsimonious way is still essential for "small T" applications. Furthermore, the presence of spatial dependence may be interesting in its own right for some applications: e.g., when one is testing for spillover effects (say, from technology, or from economic growth in neighbouring regions).

Standard tests like the Moran would reject the hypothesis of no spatial correlation in presence of either, so the appropriate modeling approach would still be ambiguous. Spatial panel estimators would also react to the presence of a factor structure indicating a significant spatial correlation SAR or SEM term, usually with a very high coefficient, even if these are actually absent. It may therefore be useful to devise a test for detecting "distance-related", or "Tobler-type" dependence even in the presence of a "factor-type" one.

2 Cross-sectional vs. spatial dependence

It is now time to try out an operational definition of what we mean by "local", "global", "spatial" dependence etc..

In the two relative strands of literature, the term *global* is used to different meanings. In recent work by Pesaran and Tosetti [4] dependence in spatial processes is characterized as being distance-decaying or not in an asymptotic fashion, introducing the concepts of cross-sectional strong dependence (CSD) and cross-sectional weak dependence (CWD). With some exceptions, the processes characterized by factor models and spatial models are shown to be typical cases of, respectively, the first and second type. To fix ideas, I shall keep the two mostly used, although less general, representations, identifying the factor model

$$y_{it} = X_{it}\beta + \gamma_i\mu_t + \epsilon_{it}$$

with CSD, and the spatial (autoregressive) error model²

$$y_{it} = X_{it}\beta + u_{it}; u_{it} = \lambda(I_T \otimes W)u_{it} + \epsilon_{it}$$

¹Unless factor loadings themselves should be made dependent on distance, of course.

²The spatial moving average model (SMA)

$$y_{it} = X_{it}\beta + u_{it}; u_{it} = \lambda(I_T \otimes W)\epsilon_{it} + \epsilon_{it}$$

would be another candidate; the dependence induced is abruptly decaying to 0 after the first order of neighbourhood, so distinctions in the following would be even sharper. In a sense, to our purpose we consider the SMA model as a "special case" of SAR, although this isn't true from an analytical viewpoint.

with CWD throughout the paper.

3 The CD and $CD(p)$ tests

Pesaran ([2]) devised a test for cross-sectional dependence in the residuals of a panel model he calls CD , which has remarkable properties in samples of any practically relevant size and is robust to a variety of settings. The test is based on averages over the time dimension of pairwise correlation coefficients between cross-sectional units:

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij} \right); \quad \hat{\rho}_{ij} = \frac{\sum_t e_{it} e_{jt}}{(\sum_t e_{it}^2)^{1/2} (\sum_t e_{jt}^2)^{1/2}}$$

The only big drawback is that the test loses power against the alternative of cross-sectional dependence if the latter is due to a factor structure with factor loadings averaging zero, that is, some units react positively to common shocks, others negatively. Apart from that, the CD test, distributed as a standard Normal, is very flexible and can be used as a formal test or as a descriptive statistic to assess the degree of cross-sectional dependence over the whole sample or over subsamples, just by reducing the set of cross-sectional units over which the test is calculated.

A variant of the CD test, called $CD(p)$ test, takes into account an appropriate subset of "neighbouring" cross-sectional units to check the null of no cross-sectional dependence against the alternative of *local* cross-sectional dependence, i.e. dependence between neighbours only. To do so, the pairs of neighbouring units are selected by means of a binary proximity matrix much alike the one used in spatial models. In the original paper, a regular ordering of observations is assumed, so that the m -th cross-sectional observation is a neighbour to the $(m-1)$ -th and to the $(m+1)$ -th and the selector matrix is bidiagonal³.

Although for now the $CD(p)$ test looks as the only viable alternative well documented in the literature, nothing would prevent us from applying the reasoning that follows to any test statistic that allows restricting its application to "neighbours". One example could be the Breusch-Pagan LM test for cross-sectional dependence and its scaled version $SCLM$ analyzed in the same paper ([2]).

3.1 The $CD(p)$ test on irregular grids

While in Pesaran's original paper ([2]) a regular ordering of observations was considered, extending the $CD(p)$ test to irregular lattices is straightforward. A binary proximity matrix is employed as a selector for discarding the correlation coefficients relative to pairs of observations that are not neighbours in computing the CD statistic. The test is defined as

$$CD = \sqrt{\frac{T}{\sum_{i=1}^{N-1} \sum_{j=i+1}^N w(p)_{ij}}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N [w(p)]_{ij} \hat{\rho}_{ij} \right)$$

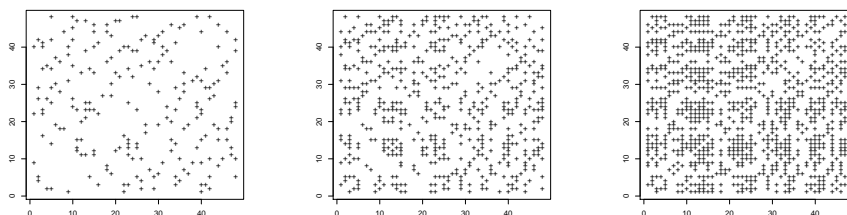
³In Pesaran and Tosetti ([4]) the first and last observations are also assumed to be neighbours, so that the "linear" world of Pesaran's paper becomes a "circular" one.

where $[w(p)]_{ij}$ is the (i, j) -th element of the p -th order proximity matrix, so that if h, k are not neighbours, $[w(p)]_{hk} = 0$ and $\hat{\rho}_{hk}$ gets "killed"; this is easily seen to reduce to formula (14) in Pesaran (cit.) for the special case considered in that paper.

3.2 Recursive $CD(p)$

The $CD(p)$ on orders greater than 1 simply uses higher-order proximity matrices, where second-order neighbours are defined as neighbours of neighbours and so on, so that an n -th order neighbour to a given region is every region that has at most $n - 1$ other regions between itself and the original one, following the "shortest" route. Of course, $(n - 1)$ -th order neighbours also qualify as n -th order ones, so increasing orders of proximity means progressively "filling" the proximity matrix, up to saturation when the order is such that every observation is the neighbour to every other one: then the only zeros left are on the main diagonal. In figure 1 we can see proximity matrices "filling up" as p increases.

Figure 1: Proximity matrices of 1st, 2nd and 3rd order, US states



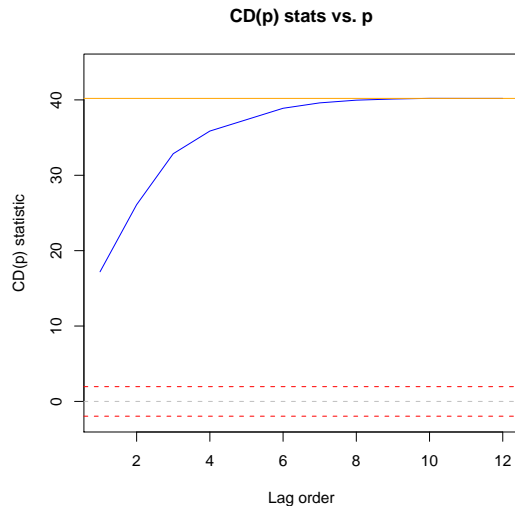
So for a given set of residuals the $CD(p)$ test must converge to CD as p increases⁴ although there is no guarantee it will do so monotonically.

As the test is distributed as standard Normal irrespective of p , critical values are the same and one might calculate it recursively on p looking for neighbourhood orders at which dependence becomes significant, and if there are more, for the value of p that maximizes the test statistic. In figure 3.2 the $CD(p)$ s for $p = 1, \dots$ until saturation are plotted for the well-known Munnell model ([1]). Dotted red lines are 5% critical values, the orange line is the CD test value.

Unfortunately, while recursive application of the $CD(p)$ test indeed looks as a useful way of heuristically assessing the pattern of cross-sectional dependence throughout the sample, the sample size on which calculation is based is clearly depending on p , and so will test power be.

⁴In the case of US 48 mainland states, the saturation point is at lag 11; for Italy's 103 provinces, it is at lag 20. Intuitively, the more the map approaches a regular grid, the lower the saturation point; the more it approaches Pesaran and Tosetti's "circular world" the nearer the saturation point is to $N/2$.

Figure 2: Recursive $CD(p)$ on Munnell’s model



3.3 The CD test under CSD and CWD

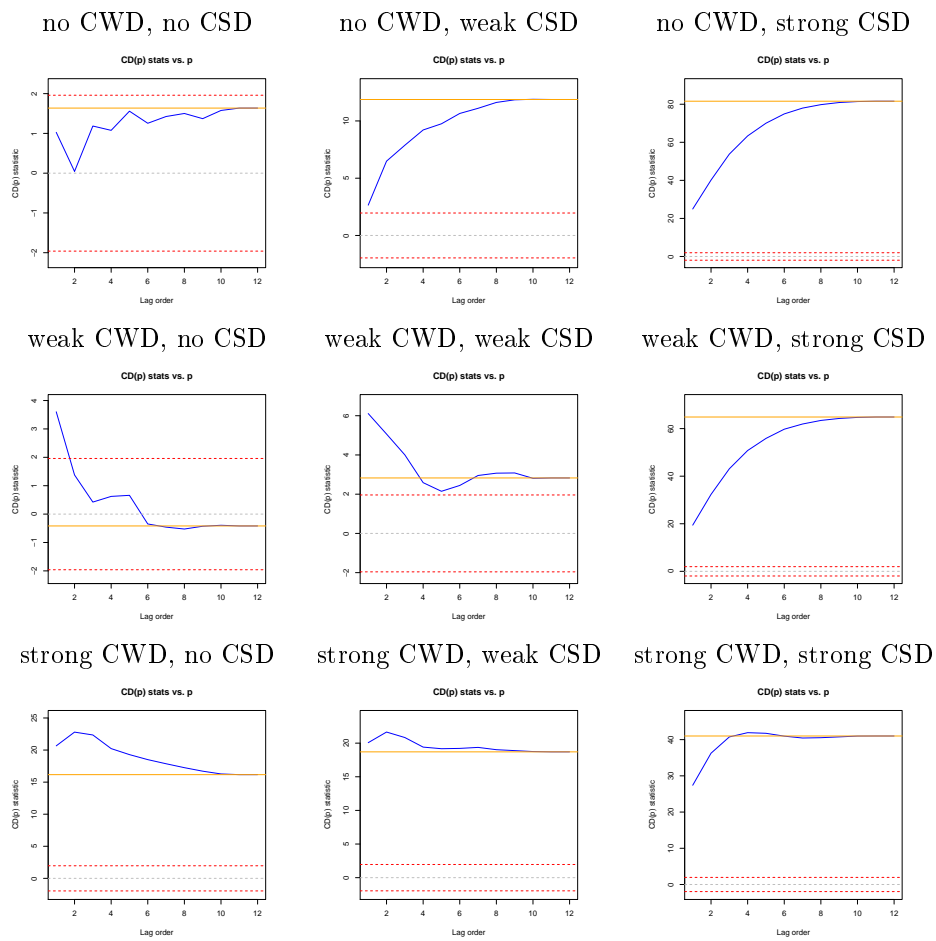
Under CSD the average correlation between units remains the same irrespective of distance. Simulations in Pesaran and Tosetti show that the $CD(p)$ test statistic has lower values than CD if dependence is indeed CSD. The reason for this is that under global dependence the $CD(p)$ turns out to be based on fewer observations (the neighbours) w.r.t. the CD , which exploits the full sample. Thus (as observed in [2]), the CD has maximum power against CSD while under CWD, on the converse, the $CD(p)$ uses fewer observations, but is likely to use those which are more strongly correlated (the neighbours), possibly offsetting the former effect. Which one will prevail will depend on the degree of spatial correlation.

In figure 3.3 some different situations are depicted, based on simulation of a model combining a simple factor structure with loadings set equal to 1 for every unit and a spatially autoregressive error term. The variance of the common factor μ_t is set, respectively, to 0, 0.3 and 1 times that of the idiosyncratic error in the no, "weak" and "strong" CSD cases. The values for λ are set to 0, 0.3 and 0.8 to exemplificate no, "weak" and "strong" CWD.

So, if we assume that there is either one or the other type of dependence, while a decreasing sequence of $CD(p)$ tests is unambiguously indicative of a CWD process, an increasing one might well be related to CWD with *local* correlation approaching unity⁵, as well as to CSD. If both CWD and CSD coexist, then the overall process is CSD (see Pesaran and Tosetti, [4]) but for the above reasons the situation as regards the test should be less clear-cut. Heuristically, a decreasing pattern for some interval should be indicative of a CWD component, but a more formal testing procedure is warranted.

⁵In the simulated example, the pattern of the $CD(p)$ statistic becomes monotonically increasing for λ between 0.93 and 0.95.

Figure 3: Different patterns for the CD(p) test



4 Testing the importance of ordering

We henceforth concentrate on testing whether there is a CWD component in the spatial process at hand, irrespective of the presence of a CSD one. In econometric terms, we test for the presence of a spatial dependence structure allowing for a common-factor one; or, if we take the encompassing model:

$$y_{it} = X_{it}\beta + \gamma_i\mu_t + u_{it}; u_{it} = \lambda(I_T \otimes W)\epsilon_{it} + \epsilon_{it}$$

as the most general structure, we test for $\lambda = 0$.

4.1 The test

While given one set of residuals the global CD statistic is uniquely determined, the local $CD(p)$ is conditional on the choice of the contiguity matrix, so that

$$CD_p^* = CD(p, W^*|\hat{e})$$

where W^* is an appropriately chosen contiguity matrix. If we consider W^* as random, then the relative $CD(p)$ statistic is itself a random number.

The testing idea is of discriminating whether, among many possible random orderings, the "true" one described by W gives rise to a significantly different $CD(p)$ statistic.

Given one set of residuals \hat{e} , one order of proximity \bar{p} and \mathscr{W} the set of all possible binary contiguity matrices W^* having the same number of ones as W ,⁶ define $CD_p(\mathscr{W}|\hat{e}, \bar{p})$ as the set of all $CD_{\bar{p}}$ tests associated with every possible draw from \mathscr{W} . Also call $CD_{\bar{p}} = CD_p(W|\hat{e}, \bar{p})$ the statistic associated to the "true" W matrix.

Intuitively, under the null of no CWD component in the process that generated the residuals, in this setting equivalent to $\lambda = 0$, the choice of W doesn't make any difference as in

$$e_{it} = \gamma_i\mu_t + B(W)^{-1}u_{it}$$

$B(W)^{-1} = (I_T \otimes (I_N - \lambda W))^{-1}$ reduces to the identity matrix for any $W \in \mathscr{W}$. Under the alternative, on the converse, the process generates errors with stronger correlation between neighbours.

So, under the null of no CWD component, $CD_{\bar{p}}$ should be interpretable as a random draw from $CD_{\bar{p}}(\mathscr{W}|\hat{e})$, while under the alternative it can be expected to take more extreme values, all this irrespective of the degree of cross-sectional correlation induced by the presence of a CSD process (here, $\gamma_i \neq 0$ for some i)⁷.

4.2 A bootstrap procedure

For all this, a non-parametric bootstrap test for spatial dependence in the error process may be based on testing the importance of ordering: if the dependence is CSD, then on average there will be no difference between the $CD(p)$ test based

⁶The number of neighbours is kept constant across the different W^* to avoid changes in test power.

⁷ $E[CD_{\bar{p}}(\mathscr{W}|\hat{e})]$, on its part, should be "far" from zero if there is a CSD component in the error process.

on the actual proximity matrix $C\bar{D}_p = CD_p(W|\hat{e}, \bar{p})$ and those based on random orderings of the observations in space, sampled from $CD_p(\mathcal{W}|\hat{e}, \bar{p})$. Thus, for any given \bar{p} , a testing procedure may be based on extracting the model's residuals, computing the "true" $CD(p)$ test and then comparing the latter with M replications of a randomized $CD(p)$ statistic obtained randomly selecting the same number of "neighbours" from all regions, which is easily accomplished by randomly sampling an appropriate number of positions to be filled with ones in the randomized W matrix. The pseudo-p value for M draws would then be

$$MCW_p = \frac{\sum_{h=1}^M I[CD_p(W_h^*|\hat{e}) \geq CD_p(W|\hat{e})]}{M}$$

where $I[.]$ is such that $I[TRUE] = 1$, $I[FALSE] = 0$.

The null of no spatial dependence would be rejected at, say, 5% significance if $MCW < 0.05$, meaning that the actual $CD(p)$ value were more extreme than the 95th quantile of the distribution of randomized values.

4.3 An illustration on Munnell's model

Considering, again, Munnell's model on public capital productivity over 48 states in the USA observed on 17 years and testing for residual cross-sectional correlation, the global CD statistic is 40.2 while the local $CD(1)$ is 17.2. Increasing the contiguity order, $CD(2) = 26.1$, $CD(3) = 32.9$ and so on, in the monotonically increasing sequence portrayed in figure 3.2.

Although Munnell's main interest was in the significance of the coefficient on public capital, one might also wonder whether the residuals from the estimated production function show evidence of spatial spillover effects and/or of common factors excluded from the model specification driving growth in all states alike (in which case inclusion of time dummies could capture the effect) or idiosyncratically, according to a factor structure.

Analyzing the recursive CD plot in figure 3.2, the most likely conclusion is that there are country-wide effects affecting every state in a non-spatial fashion, although a spatial diffusion process with coefficient very close to unity would also be a possibility. But if the first is true, as looks very reasonable in this setting where fiscal, technological or commodity-price related shocks can affect every unit irrespective of distance, is there *also* a spatial process in the errors? Applying the above test procedure we get $MCW_1 = 0.001$ *most of the time*⁸, so as to conclude against the null of no spatial process and thus in favour of spillover effects.

4.4 Montecarlo evidence on a simple model

4.4.1 Equal factor loadings

In order to have a first assessment of the test properties, although on one very particular example, I simulate a simple model with 2 stationary regressors on the same dimensions as in Munnell: 48 units and 17 time periods, according to

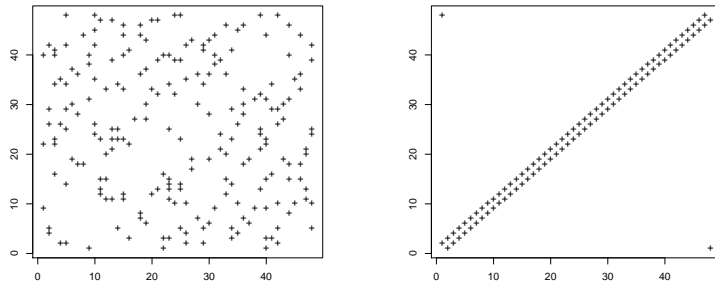
$$y_{it} = \beta_1 x_{1it} + \beta_2 x_{2it} + \mu_t + u_{it}; u_{it} = \lambda(I_T \otimes W)\epsilon_{it} + \epsilon_{it}$$

⁸Of course the quasi-p value changes every time the distribution is resampled. Some times, we get $MCW_1 = 0.002$.

i.e., the most general model in the paper with $\gamma_i = 1$ for all i , a very simple factor structure equivalent to the inclusion of time specific fixed effects, varying the ratio τ of the variance of μ_t to that of the idiosyncratic error ϵ_{it} from 0 (no CSD component) to 0.5 and 1; and a spatial autoregressive error process with λ ranging from 0 (no CWD component) to 0.8.

I employ two different proximity matrices, one from a real-world example and one from theoretical work: the "true" matrix for the US and the regular matrix as in Pesaran and Tosetti (cit.) where unit n is neighbour to $n - 1$ and to $n + 1$, and 1 and n are set as neighbours too. The matrices are depicted below in figure 4, where one can see the different degree of connectivity: the "true" matrix has 214 ones, while the "circular" one has $N * 2 = 96$.

Figure 4: "True" and "circular" proximity matrices



The results, based on 1000 replications for the test with $M = 999$ draws, are presented below in table 4.4.1 where the first row, corresponding to $\lambda = 0$, shows the empirical size of the test (which should ideally be 0.05) and the others the power of the test for different combinations of λ and τ .

Test size under validity of the null is reasonably close to 5% in the first experiment. Test power is always very good for $\lambda = 0.8$ and also for $0.8 > \lambda \geq 0.4$, at least until $\tau < 1$. Weaker spatial dependence with $\lambda = 0.2$ also gets detected reasonably well until the common shock variance gets as big as that of the idiosyncratic error.

In the experiment involving the regular proximity matrix the empirical power of the test gets better, being very good even for $\lambda = 0.2$. On the other hand, empirical size worsens, getting biased in the sense of underrejection for $\tau = 1$.

As for the main purpose, intuitively the test appears correctly to control for CSD dependence, albeit getting "blurred" as the variance of the common shocks adds to that of the idiosyncratic error.

4.4.2 Random factor loadings

A more realistic experiment involved factor loadings sampled from the Uniform distribution over the $[-1, 1]$ interval. This experiment in turn gives satisfactory results (see Table 4.4.2): although the heterogeneous factor loadings somehow "blur" the situation, the empirical size is reasonably good and the empirical

Table 1: Empirical size and power, equal factor loadings

True US			
λ	$\tau = 0$	$\tau = 0.5$	$\tau = 1$
0	0.055	0.046	0.042
0.2	0.893	0.840	0.711
0.4	1.000	1.000	0.999
0.6	1.000	1.000	1.000
0.8	1.000	1.000	1.000
Circular			
λ	$\tau = 0$	$\tau = 0.5$	$\tau = 1$
0	0.046	0.032	0.018
0.2	0.996	0.994	0.961
0.4	1.000	1.000	1.000
0.6	1.000	1.000	1.000
0.8	1.000	1.000	1.000

power always approaches one for $\lambda \geq 0.4$: substantial spatial dependence is successfully detected irrespective of the factor structure. The situation is less optimal for very weak spatial dependence ($\lambda = 0.2$): in this case, test power suffers as the variance of the common factors increases towards that of the idiosyncratic error.

Table 2: Empirical size and power, factor loadings sampled in $[-1, 1]$

True US			
λ	$\tau = 0$	$\tau = 0.5$	$\tau = 1$
0	0.039	0.050	0.074
0.2	0.900	0.728	0.281
0.4	1.000	0.999	0.972
0.6	1.000	1.000	1.000
0.8	1.000	1.000	1.000
Circular			
λ	$\tau = 0$	$\tau = 0.5$	$\tau = 1$
0	0.052	0.054	0.026
0.2	0.995	0.984	0.785
0.4	1.000	1.000	1.000
0.6	1.000	1.000	1.000
0.8	1.000	1.000	1.000

5 Conclusions

In the spatial econometrics literature, spatial error dependence is characterized by spatial autoregressive processes, which relate every observation in the cross-

section to any other with distance-decaying intensity, which is well exemplified by Tobler's First Law of Geography: "everything is related to everything else, but near things are more related than distant things". In the literature on factor models, on the converse, the degree of correlation between cross-sectional units depends only on factor loadings, so that Tobler's Law doesn't necessarily hold.

The economic meaning of the two can be much different, so that it may be useful to devise a test for detecting "distance-related" dependence in the presence of a "factor-type" one. I deal with the problem of statistically distinguishing the two forms of dependence in spatial panel data. The problem is that standard spatial correlation tests have power against both types of dependence.

My work is based on Pesaran's CD, a test for global cross-sectional dependence with good properties, and in particular on the CD(p) variant, which only takes into account p-th order neighbouring units to test for local cross-sectional dependence. While the original formulation of the test is defined with respect to a regular spatial pattern, I show an easy extension to irregular lattices which makes the test operational in applied situations.

The pattern of CD(p) as p increases can be informative about the type of dependence in the errors, but the test power changes as new pairs of observations are taken into account. I show examples of typical patterns related to combinations of different degrees of spatial and factor dependence, concluding that while the visual representation of recursive CD(p) tests is a useful descriptive tool, there are situations where the conclusions to be drawn are still ambiguous.

To resolve this ambiguity, I propose a bootstrap test based on the values taken by the CD(p) test under permutations of the neighbourhood matrix, i.e. when "resampling the neighbours". I provide Montecarlo evidence of it being able to tell the presence of spatial-type dependence in the errors of a typical spatial panel irrespective of the presence of an unobserved factor structure. The test shows correct empirical size and good power properties even for relatively low degrees of spatial dependence.

Although the simulation design is very simple and more thorough investigations have to be done as regards the empirical properties of the proposed test, the Montecarlo evidence presented in this paper is encouraging. The extension of the simulations to less simplistic data generating processes and to different sample sizes and neighbourhood structures is left for future work.

References

- [1] A. Munnell. Why has productivity growth declined? productivity and public investment. *New England Economic Review*, pages 3–22, January 1990.
- [2] M.H. Pesaran. General diagnostic tests for cross section dependence in panels. Working Paper 1229, CESifo, 2004.
- [3] M.H. Pesaran. Estimation and inference in large heterogeneous panels with a multifactor error structure. *Econometrica*, 74(4):967–1012, 2006.
- [4] M.H. Pesaran and E. Tosetti. Large panels with common factors and spatial correlation. Working Paper 2013, CESifo, 2007.