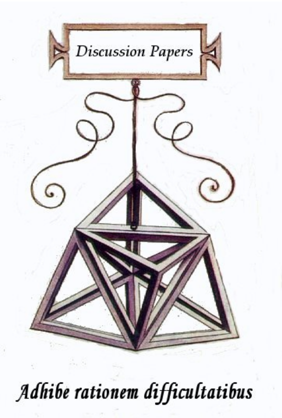
Discussion Papers
Collana di
E-papers del Dipartimento di Economia e Management – Università di Pisa



Davide Fiaschi - Lisa Gianmoena - Angela Parenti

Local Directional Moran Scatter Plot - LDMS

Discussion Paper n. 197

2015

Discussion Paper n. 197: **January 2015**

Indirizzo dell'Autore:
Dipartimenti di Economia e Managament
University of Pisa

Please use the following citation:
Fiaschi, D. L. Gianmoena, and A. Parenti (2015), "Local Directional Moran Scatter Plot - LDMS", Discussion Papers of Dipartimento di Economia e Management – University of Pisa, n. 197 (http://www.ec.unipi.it/ricerca/discussion-papers.html).

Discussion Paper
n. 197

Davide Fiaschi - Lisa Gianmoena - Angela Parenti

Local Directional Moran Scatter Plot - LDMS

Abstract

This paper propose a novel methodology to estimate the distribution dynamics of income in presence of spatial dependence by representing spatial dynamics as a random vector field in Moran space. Inference on the local spatial dynamics is discussed, including a test on the presence of local spatial dependence. The methodology also allows to compute a forecast of future income distribution which includes also the effects of spatial dependence. An application to US States is used to illustrate the effective capacities of the methodology.

Classificazione JEL: C14, O51, R11

Keywords: Exploratory data analysis, polarization, random vector field, spatial dynamics, spatial dependence, distribution dynamics, US States

Contents

I. Introduction

This paper proposes a new methodology to study the distribution dynamics of income in presence of spatial dependence.

Our proposal is the result of two converging strands of literature denoted ETSDA (Exploratory Time-Space Data Analysis) and ESTDA (Exploratory Space-Time Data Analysis) by Rey (2014).

From one hand, literature on ETSDA extends the methods used in the temporal studies on income dynamics in order to incorporate spatial dimension. Quah (1993) can be considered the pioneering contribution to ETSDA for his attempt to measure the impact of spatial dependence mapping unconditioned income levels of countries into *normalized* income levels, where normalization is respect to the incomes of neighbouring countries. Gerolimetto and Magrini (2014) represents one of more recent and most significant contribution in this line of research.

On the other hand, literature on ESTDA extends the spatial methods generally used for detecting spatial dependence in cross-sectional analysis, as the Moran's I and LISA statistics, to incorporate temporal dimension. Recently within this line of research Rey et al. (2011) have proposed the *Directional Moran Scatter Plot* to study the spatial dynamics of US states. The latter consists in analysing in the Moran space (the space defined by countries' income and its spatially lagged value) the directions of the movement vectors standardized by their beginning points, i.e. the transitions that each state has experimented between the first and the last year centered in the orgin of axes.[1]

In this paper we propose a *local* version of the Directional Moran Scatter Plot, labelled *Local Directional Moran Scatter Plot* (LDMS), which consists in the estimate of a random vector field in the Moran space exploiting the information from the observed movement vectors. With respect to Directional Moran Scatter Plot our methodology allows to conduct inference on the local spatial dynamics, and to provide a forecast of the future income distribution which takes into account also spatial dependence.

The next section gives an heuristic introduction to the Local Directional Moran Scatter Plot; Section III. discusses the nonparametric methodology used in its estimate, and how to make some inference at local level; Section IV. illustrates the use of LDMS to forecast the distribution dynamics of income in presence of spatial dependence. Section V. concludes.

II. The Local Directional Moran Scatter Plot

To gain the intuition of our proposed methodology consider Fig. 1, which reports the Moran Scatterplot (i.e. the levels of relative per capita GDP y versus

[1]The movement vectors can also be standardized by their ending point. In any case, the standardized movement vectors are placed at a common origin, but they preserve their length and direction.

its spatially lagged values Wy) for a sample of 49 US states in 1987 (red points) and 2013 (black points).[2]

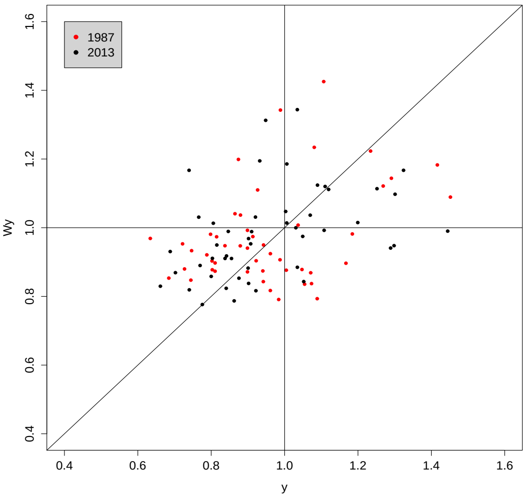


Figure 1: Moran Scatterplot of relative GDP per capita of a sample of 49 US states for 1987 (red points) and 2013 (black points). Spatial matrix W is defined by rook contiguity.

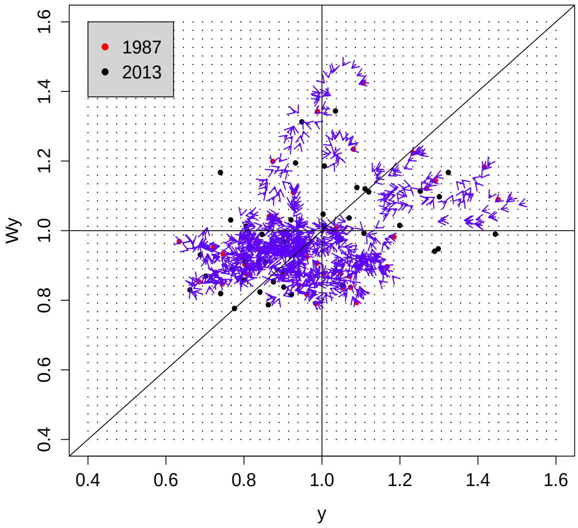


Figure 2: Movement vectors in the Moran space (y, Wy) for a sample of 49 US states over the period 1987-2013.

For both years Moran's I is positive (equal to 0.10 in 1987 and 0.13 in 2013) and statistically significant at 10%, suggesting that some spatial dependence should be at work; however, from Moran Scatter Plot no information can be extracted on the *impact* of this spatial dependence on the distribution of GDP per capita in terms of its strength and direction.[3]

A possibility to fill this gap is to assume that, in the same spirit of the distribution dynamics approach (see Quah, 1997), the dynamics of GDP per capita of an economy can be expressed as a (random) function of only its position in the Moran space, i.e. the dynamics of GDP per capita follows a Markovian process, where the states are defined in terms both of the (relative) level of GDP per capita y and its spatially lagged values Wy (instead of only y). This corresponds to the estimate of a *random vector field* in Moran space, which we label *Local Directional Moran Scatter Plot* (LDMS).[4]

[2]The spatial matrix Wy is defined by rook contiguity; given this definition of spatial dependence, we exclude from the sample the two US states without any link (Alaska and Hawaii).

[3]Using a six-nearest neighbor spatial weight matrix, Gerolimetto and Magrini (2014) finds statistically significant spatial dependence across US States. Morevoer, their estimated Moran's I is generally higher.

[4]A vector field in a plane can be visualized as a collection of arrows with a given magni-

The *movement vectors* reported in Fig. 2, representing all the observed transitions calculated with a time lag of 10 years and expressed in annual scale (1-year ahead) in the Moran space (y, Wy), provides the basic information set to estimate a LDMS.[5] In particular, Fig. 2 contains 49 x (2013-1987-10+1) = 833 movement vectors. For comparison, Rey et al. (2011) in the upper panel of their Fig. 2 reports only 49 movement vectors, representing the transitions from the first to the last year (1969 and 2008 respectively) of each state in the sample. Nonetheless the different time period considered, the overall picture of spatial dynamics looks very similar in the most of Moran space (the south-west quadrant contains the most of observations with a spatial dynamics converging toward bisector), but with some important differences (the spatial dynamics in the north-east and south-east quadrants).

Figure 2 suggests an overall pattern of convergence to bisector and, in particular, towards the region around point (1,1), although such convergence is absent or very weak for other regions of the Moran space, as for example that around (1.3,1.1). In general, the presence of a strong random component in US states' movements makes difficult to identify any spatial pattern by only a graphical inspection, especially when the number of movement vectors is very large as in our case.

In the next section we discuss how to properly estimate a LDMS by the set of observed movement vectors.

III. Estimation of a Local Directional Moran Scatterplot

Consider a sample of N economies observed for T periods; economy j is characterized by its level of relative (to the sample average) income in each point in time y_{jt}, and by the average income of its neighbours Wy_{jt}, where W is the the j-the row of the spatial weight matrix expressing which economies are neighbours of j ($j = 1, ..., N$ and $t = 1, ..., T$), and y_t is the vector of relative income of all economies.

We assume that the spatial dynamics of economy j at period t, i.e. the dynamics of economy j in the space (y, Wy), only depends on (y_{jt}, Wy_{jt}), i.e. y_{jt} follows a *time invariant* and *Markovian* stochastic process.

The spatial dynamics of the sample in the Moran space can be therefore represented by a random vector field (RVF). In particular, given a subset L of the possible realization of (y, Wy) (i.e. a lattice in Moran space, see small black points in Figure 2), a RVF is represented by a random variable $\Delta_\tau z_i$, where $\Delta_\tau z_i \equiv$

tude and direction (our movement vectors) each attached to a point in the plane (Polyanin and Manzhirov, 2006). A *random* vector field consider for each point in the plane not just a movement vector, but a set of movement vectors with an associate probability distribution (Polyanin and Manzhirov, 2006).

[5] All the calculations are made using R (R Core Team, 2014). Codes and data are available on author's web page `http://dse.ec.unipi.it/~fiaschi/`.

$(\Delta_\tau y_i, \Delta_\tau W y_i) \equiv (y_{it+\tau} - y_{it}, W y_{it+\tau} - W y_{it})$, indicating the spatial dynamics (i.e. the dynamics from period t to period $t+\tau$ represented by a movement vector) at $z_i \equiv (y_i, W y_i) \in L$.

For each point in the lattice z_i, with $i = 1, ..., L$, we therefore estimate the distribution of probability $\Pr(\Delta_\tau z | z_i)$ on the $N(T-\tau)$ observed movement vectors $\Delta_\tau^{OBS} z$. In particular, $\Pr\left(\Delta_\tau^{OBS} z_{jt} | z_i\right)$ measures the probability that the dynamics at z_i follows $\Delta_\tau^{OBS} z_{jt}$; this suggests that $\Pr\left(\Delta_\tau^{OBS} z_{jt} | z_i\right)$ should decrease as function of the distance between z_i and z_{jt}^{OBS}.

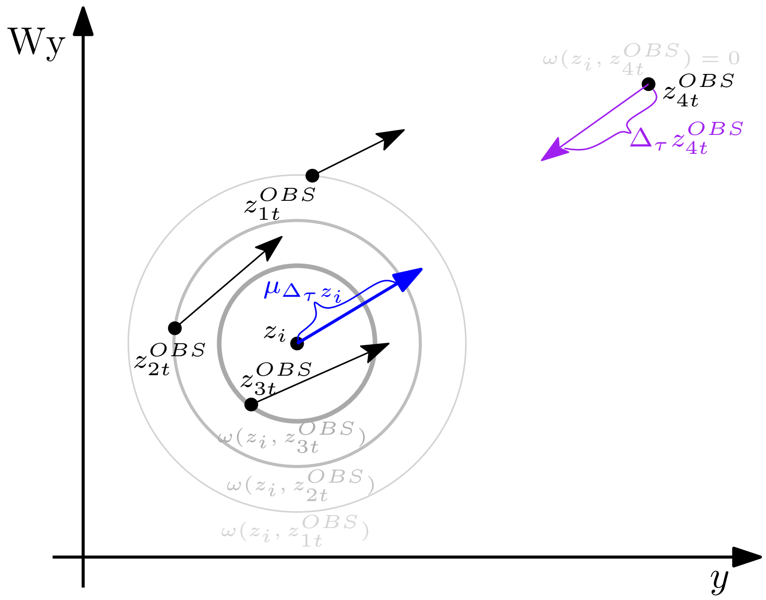


Figure 3: Local mean estimation of the *expected movement* from z_i ($\widehat{\mu}_{\Delta_\tau z_i}$) from four observed movement vectors (z_{jt}^{OBS}). Probabilities attached to each observed movement vectors, given by $\omega\left(z_i, z_{jt}^{OBS}\right)$, are a negative function of the distance between z_i and z_{jt}^{OBS}.

Following this intuition Fig. 3 depicts a point of the lattice z_i and four observed movement vectors, which origin at different distance from z_i. Function $\omega\left(z_i, z_{jt}^{OBS}\right)$ measures for each observed movement vector its probability to affect the movement at z_i; these probabilities decline with distance from z_i (i.e. $\omega\left(z_i, z_{1t}^{OBS}\right) < \omega\left(z_i, z_{2t}^{OBS}\right) < \omega\left(z_i, z_{3t}^{OBS}\right)$), and very far observed movement vectors should have zero probability ($\omega\left(z_i, z_{4t}^{OBS}\right) = 0$). Blue vector is the *expected movement* from z_i, $\mu_{\Delta_\tau z_i}$, calculated on the base of the distribution of probabilities on the observed movement vectors.

A convenient way to calculate these probabilities is to use a kernel function to measure the distance between z_i and z_{jt}^{OBS}. In particular:

$$\omega\left(z_i, z_{jt}^{OBS}\right) = \frac{K\left(\frac{(z_i - z_{jt}^{OBS})^T S^{-1}(z_i - z_{jt}^{OBS})}{h^2}\right)\frac{\det(S)^{-\frac{1}{2}}}{2h^2}}{\sum_{t=1}^{T-\tau}\sum_{j=1}^{N} K\left(\frac{(z_i - z_{jt}^{OBS})^T S^{-1}(z_i - z_{jt}^{OBS})}{h^2}\right)\frac{\det(S)^{-\frac{1}{2}}}{2h^2}} \quad (1)$$

is assumed to be an estimate of the probability that at z_i spatial dynamics follows observed movement vectors $\Delta_\tau^{OBS} z_{jt}$, where $K(\cdot)$ is the kernel function, h is the smoothing parameter and S is the sample covariance matrix of z^{OBS}. The kernel function $K(\cdot)$ is generally a smooth positive function which peaks at 0 and decreases monotonically as the distance between the observation z_{jt} and the point of interest z_i increases (see Silverman, 1986 for technical details). The smoothing parameter h controls the width of the kernel function.[6] In the estimation we use a multivariate Epanechnikov kernel (see Silverman, 1986, pp. 76-78), i.e.:

$$K(u^T S^{-1} u) = \begin{cases} \frac{2}{\pi}(1 - u^T S^{-1} u) & \text{if } u^T S^{-1} u < 1 \\ 0 & \text{if } u^T S^{-1} u \geq 1, \end{cases} \quad (2)$$

where $u \equiv \left(z_i - z_{jt}^{OBS}\right)/h$. Multivariate Epanechnikov kernel is particularly adapted to our scope because it assigns zero probability to observed movement vectors very far from z_i.[7] The exact quantification of "very far" is provided by bandwidth h, i.e. higher bandwidth means higher number of observed movement vectors entering in the calculation of the movement at z_i.

Given Eq. (1) for each point in the lattice z_i we estimate the τ-period ahead *expected movement* $\mu_{\Delta_\tau z_i} \equiv E\left[\Delta_\tau z_i | z_i\right]$ using a *local mean estimator*, firstly proposed by Nadaraya (1964) and Watson (1964), where the observations are weighted by the probabilities derived from the kernel function, i.e.:[8]

$$\widehat{\mu}_{\Delta_\tau z_i} = \sum_{t=1}^{T-\tau}\sum_{j=1}^{N}\omega\left(z_i, z_{jt}^{OBS}\right)\Delta_\tau z_{jt}^{OBS} = \Pr\widehat{(\Delta_\tau z | z_i)}\Delta_\tau z^{OBS}. \quad (3)$$

The estimation of Eq. (3) strongly depends on the choice of τ. This choice is the result of a trade-off: from one hand, a too short τ can increase the noise in the estimation due to the possible presence of business-cycle fluctuations; on the other hand, a too long τ could contrast with the local characteristics of the estimate, increasing the probability that observed movement vectors very far from z_i affects the estimate of $\mu_{\Delta_\tau z_i}$.[9]

[6] In all the estimation we use the optimal normal bandwidth; for a discuss on the choice of bandwidth see Silverman (1986).

[7] Other possible kernels, as the Gaussian, does not allow such possibility.

[8] See Bowman and Azzalini (1997) for details.

[9] For samples with a very short time span a further limit to the choice of a long τ is the relatively strong loss of observations.

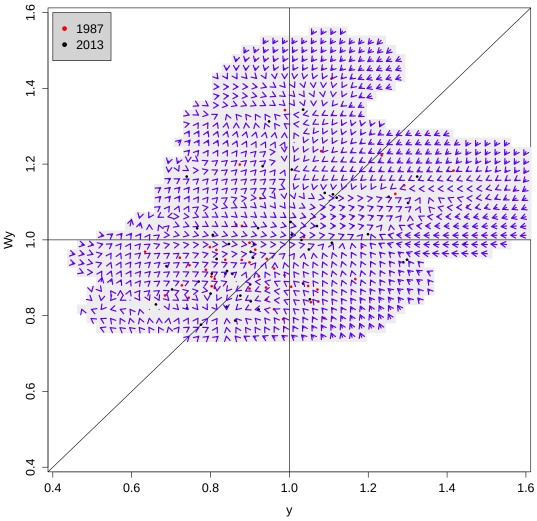


Figure 4: The Local Directional Moran Scatter Plot including all the annualized 10-years ahead expected movements for the points in the lattice where observed movement vectors are available.

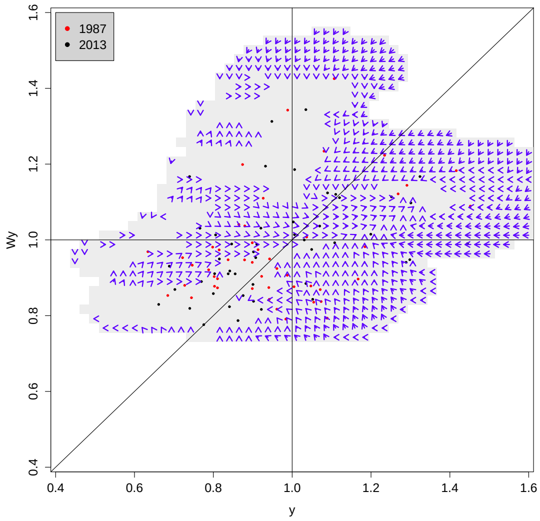


Figure 5: The Local Directional Moran Scatter Plot including all the annualized 10-years ahead expected movements for the points in the lattice where the estimated movements are statistically significant at 5% level.

Figure 4 reports the *annualized* 10-year ahead expected movements based on Eq. (3) for a lattice $50 \times 50 = 2500$ points in the range $(0.4-1.6)\times(0.4-1.6)$ in Moran space. For a wide area of Moran space we cannot calculate any expected movement due to lack of observed movement vectors sufficiently close to the points in the lattice (as discussed above such threshold in the distance is proportional to the bandwidth h).

The overall spatial dynamics pattern suggested by the estimated expected movements in Figure 4, is convergence toward a region around the bisector close but below point (1,1) for the most of trajectories starting from points in the lattice below the horizontal line $Wy = 1$; while for points above $Wy = 1$ a convergence toward regions around points (1.3,1.1) and (1,1.3) is expected.

These findings are confirmed also after controlling for the statistical significance of the estimated expected movements, whose results are reported in Figure 5 (see the next section for the bootstrap procedure used for the inference). In particular, the overall convergent spatial dynamics is confirmed; in addition, within regions previously identified as loci of convergence no significant spatial dynamics is present, suggesting that they are indeed regions of steadiness.

With respect to Rey et al. (2011) our analysis confirms the presence of spatial dependence, i.e. y and Wy tends to have the same sign of variation over time (the movement vectors show an orientation from south-west to north-east or vice versa); however, the proposed *Standardized Directional Moran Scatter Plot* reported in the bottom panel of Fig. 2 in Rey et al. (2011) cannot identify the remarkable heterogeneity of spatial dynamics among different regions of Moran space emerging from Fig. 5, and its implications in term of the existence of regions of steadiness.

III.A. Inference on Local Directional Moran Scatter Plot

Below we discuss in details how we have conducted the inference on the estimated expected movements by a bootstrap procedure, whose results is reported in Fig. 5.

Given the observed sample of observations z_{jt}^{OBS}, with $j = 1, ..., N$ and $t = 1, ..., T$, the bootstrap procedure consists of four steps.

1. Estimate the expected value of the τ-period ahead movement $\mu_{\Delta_\tau z_i}$ by Eq. (3) for each point of the lattice ($i = 1, ..., L$).

2. Draw B samples $z^b = \left(z_1^b, ..., z_{N(T-\tau)}^b\right)$ and the associated $\Delta_\tau^b z = \left(\Delta z_1^b, ..., \Delta z_{N(T-\tau)}^b\right)$, with $b = 1, ..., B$, by sampling with replacement from the observed z^{OBS} and the associated movement vectors $\Delta^{OBS} z$.

3. For every bootstrapped sample b and for each point of the lattice i estimate by Eq. (3) the expected value of the τ-period ahead movement $\mu_{\Delta_\tau z_i}^b$.

4. Calculate the two-side p-value of the estimated movement vector at point i in the lattice under the null hypothesis of no dynamics (note that null hypothesis of no dynamics is separately tested in the two directions y and Wy) as:

$$\widehat{ASL}_i = 2 \times \min\left(\sum_{b=1}^{B} \hat{\mu}_{\Delta_\tau z_i}^b \leq 0, \sum_{b=1}^{B} \hat{\mu}_{\Delta_\tau z_i}^b > 0\right) / B. \tag{4}$$

In the analysis we have set $B = 300$, and used the usual significance level of 5% to decide which expected movements to report in Fig. 5.

III.B. Test on the Presence of Local Spatial Dependence

The local characteristics of LDMS allows also to test on the presence of *local* spatial dependence in the same spirit of LISA (see Anselin, 1995). In particular, the null hypothesis of no spatial dependence in the estimated movements in Moran space can be formulated as follows:

$$H_0 : \mu_{\Delta_\tau z_i} \equiv E\left[\Delta_\tau z_i | z_i\right] = E\left[\Delta_\tau z_i | y_i\right], \tag{5}$$

that is the null hypothesis is that the dynamics in point z_i only depends on the value of y_i.

In the following we describe a permutation test to test the null hypothesis in Eq. (5).

1. Generate P independent permutation samples $z^p = (y^p, Wy^p)$, with $p = 1, ..., P$, by taking the entire time-series of an economy but randomly permuting its neighbours (therefore in every permutation sample y^p is equal to y^{OBS}, but Wy^p will be randomly different from Wy^{OBS}).

2. For every permutation sample p and for each point of the lattice i by Eq. (3) estimate the expected value of the τ-period ahead movement $\mu^p_{\Delta_\tau z_i}$.

3. For each point of the lattice i calculate the difference $\Delta_{\hat{\mu}_i} \equiv \hat{\mu}_{\Delta_\tau z_i} - \hat{\mu}^p_{\Delta_\tau z_i}$.

4. Calculate the two-side p-value of the estimated movement vector at point i in the lattice under the null hypothesis of no spatial dependence as

$$\widehat{ASL}_i^P = 2 \times \min \left(\sum_{p=1}^{P} \Delta_{\hat{\mu}_i} \leq 0, \sum_{p=1}^{P} \Delta_{\hat{\mu}_i} > 0 \right) / P. \tag{6}$$

Figure 6, which reports the results of the permutation test for $P = 300$ permutations, highlights how within the three regions previously indicated as of steadiness, spatial dependence is absent, while is particularly effective at the borders of the north-east quadrant. Spatial dependence appears to be a significant force also around (1,1).

The overall picture suggests that spatial dependence is a pervasive phenomenon, but its effects appears not so important at aggregate level because it is not significant in the regions of steadiness, where the most of US states are concentrated.[10]

IV. Forecasting by a Local Directional Moran Scatter Plot

The estimated LDMS also allows to compute $F \times \tau$-year ahead projections starting from the observed cross-economy income distribution. The proposed procedure is similar in the distribution approach to the use of the estimated stochastic kernel to project in the future the actual distribution; in the limit such projection leads to the ergodic (equilibrium) distribution. In particular, the randomness of the estimated LDMS suggests to replicate S time the procedure of computation of the F-period ahead projection and to calculate the average distribution (the replications allows also to calculate confidence bands for our $F \times \tau$-year ahead projected distribution).

The procedure for the computation of the $F \times \tau$-year ahead distribution starting from the distribution in the last year T z_T^{OBS} is as follows.

[10]From the point of view of distribution dynamics the regions of steadiness can be seen as also the loci in the plane where the US states should pass more time.

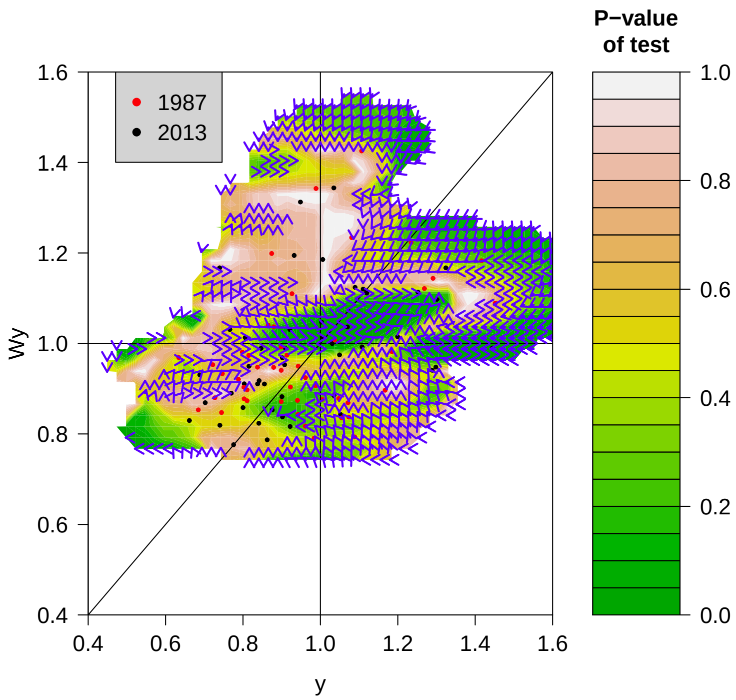


Figure 6: Permutation test for the presence of local spatial dependence in RVF

1. For each replication s, with $s = 1, ..., S$:

   (a) For each economy $j = 1, ..., N$ set $z_j^f = z_j^{f-1}$ $(z_j^1 = z_{jT}^{OBS})$.

   (b) For each economy j individuate the closest point in the lattice, and assign to the economy j the estimated probability distribution on the observed movement vectors for that point, i.e. $\widehat{\Pr\left(\Delta_\tau z | z_{i_j^*}\right)}$, where $i_j^* = \text{argmin}_{\{i\}_{i=1}^L} \parallel z_j^f - z_i \parallel$.

   (c) For each economy j draw one transition, denoted by $\Delta_\tau z_j^f$, from the observed $\Delta_\tau^{OBS} z$ with probability $\widehat{\Pr\left(\Delta_\tau z | z_{i_j^*}\right)}$ and calculate $\tilde{z}_j^{f+1} = z_j^f + \Delta_\tau^f z_j^f$.

   (d) For each economy j normalize $y_j^{f+1} = \frac{\tilde{y}_j^{f+1}}{\sum_{j=1}^N \tilde{y}_j^{f+1}}$ in order to maintain that the new calculated distribution y_j^{f+1} has mean one.

   (e) For each economy j calculate the spatial lagged value Wy_j^{f+1} and set $z_j^{f+1} = \left(y_j^{f+1}, Wy_j^{f+1}\right)$.

(f) Repeat steps (a)-(e) for $f = 2, ..., F$.

(g) Estimate the cross-economy income distribution for the last forecast period F $\hat{\phi}^s_{F\times\tau} = \hat{f}\left(y^F\right)$.

2. Take the average of the estimated income distributions on all replications $\bar{\hat{\phi}}_{F\times\tau} = \sum_{s=1}^{S} \hat{\phi}^s_{F\times\tau}/S$ as the expected $(F \times \tau)$-year ahead forecast distribution.

Setting $F = 5$ and $S = 1000$, the distribution in Moran space of all computed 50-year ahead forecasts for a total of 49.000 points (49 US states times 1000 replications) reported in Fig. 7 highlights how no particular dispersion/polarization emerges from the computation of forecast distributions; the most of economies is expected to populate the regions of steadiness, and a region in the north-east quadrant just above horizontal line $Wy = 1$.

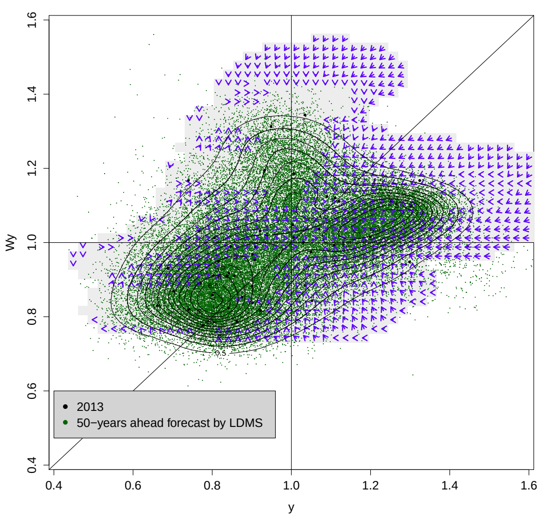


Figure 7: All computed 50-year ahead forecasts for a total of 49.000 points (49 US states times 1000 replications) (green points), the observed US states in 2013 (black points), and the estimated LDMS.

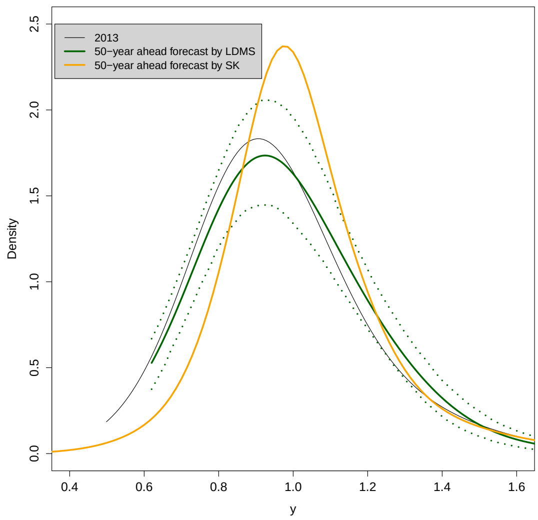


Figure 8: Expected 50-year ahead forecast cross-economy income distribution (green line), its 95% confidence bands (green dotted lines), the observed distribution in 2013 (black line) and the 50-year ahead forecast distribution calculated by the estimated stochastic kernel (SK) (orange line).

Fig. 8 shows how the computed 50-year ahead forecast distribution of y (green line) is not statistically different from the estimated income distribution in 2013 (black line). The actual income distribution across US states therefore should tend to persist at least for the next 50 years.

We also report the 50-year forecast distribution of y based on the use of a stochastic kernel (SK) (orange line), which appears centered around 1 and more or less symmetric.[11] The difference between the 50-year ahead forecast distribution calculated by LDMS and SK measures the magnitude of the bias in the estimate of distribution dynamics deriving from the omission of spatial dependence.

With a very different approach Gerolimetto and Magrini (2014) find a similar results. By using quarterly data for 48 conterminous US states over three decades running between 1981:Q1 and 2010:Q4, they show that neglecting spatial dependence substantially affect the estimate of distributional tendencies; in particular, in the second decade (1991:Q1 and 2000:Q4) the spatial estimator shows a stronger tendency towards divergence in the ergodic distribution with respect to the non spatial estimator.[12]

V. Concluding Remarks

This paper has proposed a novel methodology to analyse the distribution dynamics in presence of spatial dependence by estimating a random vector field in Moran space. The methodology has successfully identified local heterogeneity in spatial dynamics for US States from 1987 to 2013. Inference on such local heterogeneity has shown how spatial dependence is present only in some regions of Moran space, and that there exists a converging dynamics to three regions where local spatial dependence is instead very weak. The forecast of future income distribution suggests that the most of US States should persist within the three regions, and that no particular change is expected in the income distribution with respect to 2013. The comparison with the forecasted distribution calculated by stochastic kernel generally used in the distribution dynamics literature has shown how the former can be bias from the omission of spatial dependence.

The methodology could be refined by adopting an adaptive kernel in the estimation of LDMS, i.e. a kernel whose bandwidth changes accordingly to the density of observation around the point in the lattice (in particular, the bandwidth is larger where observations are less numerous, see Silverman, 1986). More important, the analysis can be extended to include other explanatory variables of the movement vectors (e.g. the typically Solovian variables such as investment rates and population growth); the limit is the so-called "curse of dimensionality" that generally plagues the use of kernel in multivariate analysis (see, again, Silverman, 1986).

[11] We use a Gaussian kernel with adaptive bandwidth. See Quah (1997) for more details on the meaning of stochastic kernel and its use to forecast future distributions, and (Silverman, 1986, p. 100) for the procedure to estimate a stochastic kernel with adaptive bandwidth.

[12] Their results are not exactly comparable with ours both for the data used in the analysis (we have a more limited time span), and for the estimation of the stochastic kernel made using a nearest-neighbor bandwidth in the first year, normal scale bandwidth in the last year, a Gaussian kernel and mean bias adjustment via a local linear estimate.

Acknowledgements We thank all the participants of the 13*rd* International Workshop Spatial Econometrics and Statistics in Toulon, April 2014 for their helpful comments. Usual disclaimer applies.

Affiliation and Address of the Authors

Davide Fiaschi: Dipartimento di Economia e Management, University of Pisa, Pisa, Italy.

Lisa Gianmoena: Institute of Advanced Studies (IMT), Lucca, Italy.

Angela Parenti:Institute of Advanced Studies (IMT), Lucca, Italy.

References

Anselin, L. (1995). Local indicators of spatial association - LISA. *Geographical Analysis 27*(2), 93–115.

Bowman, A. and A. Azzalini (1997). *Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations*. Oxford: Oxford University Press.

Gerolimetto, M. and S. Magrini (2014). Spatial distribution dynamics. mimeo.

Nadaraya, E. (1964). On estimating regression. *Theory of Probability & Its Applications 9*(1), 141–142.

Polyanin, A. and A. Manzhirov (2006). *Handbook of mathematics for engineers and scientists*. CRC Press.

Quah, D. (1993). Empirical cross-section dynamics in economic growth. *European Economic Review 37*(2-3), 426–434.

Quah, D. (1997). Empirics for Growth and Distribution: Stratification, Polarization, and Convergence Clubs. *Journal of Economic Growth 2*(1), 27–59.

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rey, S. (2014). Spatial dynamics and space-time data analysis. In *Handbook of Regional Science*, pp. 1365–1383. Springer.

Rey, S., A. Murray, and L. Anselin (2011). Visualizing regional income distribution dynamics. *Letters in Spatial and Resource Sciences 4*(1), 81–90.

Silverman, B. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Watson, G. (1964). Smooth regression analysis. *Sankhyā: The Indian Journal of Statistics, Series A*, 359–372.

Discussion Papers
Collana del Dipartimento di Economia e Management, Università di Pisa (ISSN: 2039-1854)