CONTRIBUTI DI RICERCA CRENOS

A FLEXIBLE SPECIFICATION OF SPACE–TIME AUTOREGRESSIVE MODELS

Edoardo Otranto
Massimo Mucciardi

WORKING PAPERS

2016/08

CUEC

CENTRO RICERCHE ECONOMICHE NORD SUD
(CRENoS)
UNIVERSITÀ DI CAGLIARI
UNIVERSITÀ DI SASSARI

CRENOS was set up in 1993 with the purpose of organising the joint research effort of economists from the two Sardinian universities (Cagliari and Sassari) investigating dualism at the international and regional level. CRENoS' primary aim is to improve knowledge on the economic gap between areas and to provide useful information for policy intervention. Particular attention is paid to the role of institutions, technological progress and diffusion of innovation in the process of convergence or divergence between economic areas. To carry out its research, CRENoS collaborates with research centres and universities at both national and international level. The centre is also active in the field of scientific dissemination, organizing conferences and workshops along with other activities such as seminars and summer schools.
CRENoS creates and manages several databases of various socio-economic variables on Italy and Sardinia. At the local level, CRENoS promotes and participates to projects impacting on the most relevant issues in the Sardinian economy, such as tourism, environment, transports and macroeconomic forecasts.

www.crenos.it
info@crenos.it

CRENoS – CAGLIARI
VIA SAN GIORGIO 12, I-09100 CAGLIARI, ITALIA
TEL. +39-070-6756406; FAX +39-070- 6756402

CRENoS - SASSARI
VIA MURONI 23, I-07100 SASSARI, ITALIA
TEL. +39-079-213511

Title: A FLEXIBLE SPECIFICATION OF SPACE–TIME AUTOREGRESSIVE MODELS

ISBN: 978 88 9386 001 7

First Edition: June 2016
First Reprint: September 2016

by Sardegna Novamedia Soc. Coop.
Via Basilicata n. 57/59 - 09127 Cagliari
Tel. e Fax +39070271573

A Flexible Specification of Space–Time AutoRegressive Models

Edoardo Otranto
Department of Economics and CRENoS,
University of Messina, Piazza Pugliatti, 1 - 98122 Messina, Italy (*eotranto@unime.it*)
Massimo Mucciardi
Department of Economics,
University of Messina, Piazza Pugliatti, 1 - 98122 Messina, Italy (*mucciard@unime.it*)

Abstract

The Space–Time Autoregressive (STAR) model is one of the most widely used models to represent the dynamics of a certain variable recorded at several locations at the same time, capturing both their temporal and spatial relationships. Its advantages are often discussed in terms of parsimony with respect to space-time VAR structures because it considers a single coefficient for each time and spatial lag for the full time span and the full location set. This hypothesis can be very strong; the presence of groups of locations with similar dynamics makes it more realistic. In this work we add a certain degree of flexibility to the STAR model, providing the possibility for coefficients to vary in groups of locations, proposing a new class of flexible STAR models. Such groups are detected by means of a clustering algorithm. The new class or model is compared to the classical STAR and the space-time VAR by simulation experiments and a practical application.

Keywords: clustering; forecasting; space–time models; spatial weight matrix.

JEL Classification: C3, C4, C5

1 Introduction

Since the mid-seventies the research in statistical models describing the space–time evolution of real series has been widely diffused with several methodological contributions and applications, devoted to capturing both the dynamics along time and the correlations based on spatial relationships. After the seminal paper of Cliff and Ord (1975), the space-time models were extended from Pfeifer and Deutsch (1980), who propose the Space–Time ARIMA (STARIMA) class of models, an extension of the ARIMA class developed for time series to include the linear dependencies in both space and time dimensions; spatial dependencies are imposed by means of a spatial weight matrix, which incorporates spatial features such as distances between locations and neighboring sites. The consideration of the spatial structure in economic, social and environmental data sets is present in several papers; excellent reviews can be found in Anselin (1988) and Haining (1990).

The importance of considering the presence of spatial dependencies in the forecasting performance of the models was verified by Giacomini and Granger (2004) and Arbia et al. (2011). They utilized Monte Carlo experiments to show that the use of separate univariate forecasts for each region, ignoring spatial dependence (in other words, considering only the time dimension), leads to highly inaccurate forecasts. They used the most widespread space-time model, the so–called STAR(1,1), where the same coefficients, referred to the time and spatial lag, are used for the full time span and the full location set.

The success of the STAR model is due to its simple (linear) form and the possibility of including the effects of spatial autocorrelation in forecasting, because the spatial effects are considered with a time lag, differently from the purely Spatial AR (SAR; Besag, 1974), where the spatial relationships are considered only simultaneously. Moreover, LeSage and Pace (2009) showed that the Space–Time AutoRegressive (STAR) model implies a long–run steady–state equilibrium model equivalent to the SAR model.

The STAR(1,1) model is a very parsimonious representation of space–time series, but imposes strong constraints in the behavior of the spatial units over time. An unconstrained model would consider different spatial and time coefficients for each spatial unit (call it Unconstrained STAR–USTAR model); as the spatial dimension increases, the estimation becomes unfeasible, incurring in the so–called curse of dimensionality, which causes inaccuracy and uncertainty in the estimation of the model (see Giacomini and Granger, 2004). A good compromise would be a more flexible STAR(1,1) model, where the coefficients change only for groups of spatial units and not for each spatial unit, as in the USTAR model. The groups could be detected on the basis of information about the similarity of the locations (for example, geographical aggregations), but this might be subjective. We propose a procedure to detect these groups, based on a clustering agglomerative algorithm, which has some similarity to the method developed by Otranto (2010) to detect financial assets with similar conditional dynamic correlation structure. The performance of this procedure is evaluated in terms of simulation experiments, using different time spans and number of locations, considering uncorrelated and correlated disturbances and different spatial weights. Also, we apply this procedure to a real set of data, comparing the in–sample and out–of–sample forecasting performance of our model (called the Flexible STAR–FSTAR) with that of STAR and USTAR models; in this case the purpose is to obtain results showing a significant improvement of the FSTAR model performance with respect to the STAR model and a not significantly worse performance than the USTAR model.

The paper is organized as follows: section 2 recalls the STAR and USTAR models and discusses the new FSTAR model; section 3 provides details on the algorithm for the identification of the locations with similar STAR structure; section 4 discusses the Monte Carlo experiments to assess the behavior of the FSTAR model; section 5 applies and compares the three models to demographic Italian regional data. Some final remarks complete the article.

2 The Flexible STAR Model

Let us consider a set of space–time observations, relative to N locations at T different times, collected in a matrix $\boldsymbol{Y} = \{y_{i,t}\}$ ($i = 1 \ldots, N$ and $t = 1, \ldots, T$).

The classical STAR(1,1) model follows a particular autoregressive dynamics with one time lag and one spatial lag:

$$y_{i,t} = \phi y_{i,t-1} + \psi \sum_{j=1}^{N} w_{ij} y_{j,t-1} + \varepsilon_{i,t} \tag{2.1}$$

where ϕ represents the coefficient of the time–lagged effect, whereas ψ is the coefficient of the spatial lagged effect.[1] The number w_{ij} is the $ij-$th element of a $N \times N$ weight matrix $\boldsymbol{W}$ representing the spatial link between location i and location j ($w_{ij} = 0$ when $i = j$, nonnegative otherwise); the matrix $\boldsymbol{W}$ is normalized to have each row summing up to one. This matrix is generally fixed a priori and reflects the geographical characteristics of the spatial locations in terms of neighboring, distance, etc.; anyway it can represent other characteristics, such as economic distance (see. for example, Pinkse and Slade, 1998, Otranto et al., 2016). It is evident that, fixing $\boldsymbol{W}$ in advance, the estimation of the model will be subject to possible misspecification of the weight matrix (see Stetzer, 1982). For this reason, in our successive applications, we will experiment several weight matrices, verifying the robustness of our results. A necessary, but not sufficient, condition is given by $|\phi + \psi| < 1$ (see, for example, Arbia et al., 2011).

The variable $\varepsilon_{i,t}$ is a zero mean white noise, very often considered uncorrelated across regions. It is possible to relax this hypothesis in several ways; a common specification considers the presence of correlation in the spatial dimension with the following parameterization (see Anselin, 1988):

$$\boldsymbol{\varepsilon}_t = \rho \boldsymbol{W} \boldsymbol{\varepsilon}_t \tag{2.2}$$

where $\boldsymbol{\varepsilon}_t$ is a vector containing the disturbances relative to the N locations at time t and ρ is a scalar ranging in $[-1; 1]$.

Collecting the spatial observations at time t in a $(N \times 1)$ vector $\boldsymbol{y}_t$, model (2.1) can be expressed in a VAR form (see Lütkepohl, 1993):

$$\boldsymbol{y}_t = \boldsymbol{A} \boldsymbol{y}_{t-1} + \boldsymbol{\varepsilon}_t \tag{2.3}$$

where $\boldsymbol{A} = \{a_{ij}\}$ and $a_{ij} = \phi + \psi w_{ij}$.

Model (2.1) is very parsimonious, adopting the same pair of coefficients (ϕ and ψ) for all the locations and time periods, but it has good forecasting properties, as shown in Giacomini and Granger (2004) and Arbia et al. (2011).

The basic assumption of the STAR model is that each observation y_{it} depends on the past observation of the same variable in the same location i and in the neighboring locations, and this dependence is constant across the locations and over time. This strong assumption could be too restrictive, particularly in terms of fitting of the STAR(1,1) model.

[1] Assuming that $\psi_{ij} = \psi w_{ij}$ the process is isotropic (Arbia et al., 2011).

A more flexible model could be obtained allowing for changing in the parameter values in the correspondence of different locations. In a vectorial form, this model (call it Unconstrained STAR–USTAR) would assume the form:

$$\boldsymbol{y}_t = \boldsymbol{\phi} \odot \boldsymbol{y}_{t-1} + \boldsymbol{\psi} \odot (\boldsymbol{W}\boldsymbol{y}_{t-1}) + \boldsymbol{\varepsilon}_t \tag{2.4}$$

where $\odot$ is the Hadamard product, $\boldsymbol{\phi}' = (\phi_1, \ldots, \phi_N)'$ and $\boldsymbol{\psi}' = (\psi_1, \ldots, \psi_N)'$, with $|\phi_i + \psi_i| < 1$ (for $i = 1, \ldots, N$). In terms of the VAR representation (2.3), the elements of the matrix of autoregressive coefficients will be $a_{ij} = \phi_i + \psi_i w_{ij}$.

Model (2.4) is clearly more flexible but, at the same time, requires a large number of coefficients ($2(n-1)$ more than model (2.1)), involving a curse of dimensionality problem. This problem could be avoided if we were able to identify a small number of groups of locations following the same STAR dynamics. In practical terms, the idea of a parsimonious but flexible STAR(1,1) model is to detect G groups of locations so that the vector of coefficients $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ assume the form:

$$\boldsymbol{\phi} = \begin{bmatrix} \phi_1 \boldsymbol{\iota}_{n_1} \\ \vdots \\ \phi_G \boldsymbol{\iota}_{n_G} \end{bmatrix}; \qquad \boldsymbol{\psi} = \begin{bmatrix} \psi_1 \boldsymbol{\iota}_{n_1} \\ \vdots \\ \psi_k \boldsymbol{\iota}_{n_G} \end{bmatrix} \tag{2.5}$$

where n_k ($k = 1, \ldots, G$) is the size of each group ($\sum_{k=1}^{G} n_k = N$). We call this model the FSTAR(1,1) model, where F stands for Flexible.

A crucial problem is the correct assignment of each location to each group; this could be made with some a priori choice, subjectively fixing the locations in each group. This choice is more difficult with respect to the subjective choice of the weight matrix $\boldsymbol{W}$ because for the latter we can use objective information, such as the geographical distance. It is relative only to the spatial dimension; in (2.5) the G groups of coefficients are relative to both the spatial and time dimensions.

A more suitable choice could be based on some agglomerative algorithm to provide a clustering of locations; this idea is connected to the time series clustering because, according to the VAR representation (2.3), each location is characterized by a certain time series following a certain dynamic. We propose a test-based agglomerative algorithm possessing this characteristic which will be detailed in the next section.

3 Identifying Locations with Similar STAR Structure

The VAR representation (2.3) qualifies our research to be added to the time series clustering literature (for an extensive review, see Liao, 2005). In fact our purpose is similar to the problems managed, for example, in Maharaj (1999) and Otranto (2008, 2010), where a set of time series is classified according to the similarity of the autoregressive parameters characterizing the underlying processes, by the means of a hierarchical clustering algorithm based on a Wald test statistic.

In this space–time framework, the purpose is to insert the locations with similar coefficients ϕ and ψ in the same group. The algorithm we propose can be synthesized in the following steps:

1. estimate all the N univariate models and put $N^* = N$ and $G = 0$;
2. verify the joint null hypothesis $\phi_h = \phi_i$ and $\psi_h = \psi_i$ for each h and i and, if at least one hypothesis is not rejected, select the two series with maximum p-value and put $G = G + 1$ and $N^* = N^* - 2$; otherwise stop the algorithm and put $G = G + N^*$;
3. estimate the STAR(1,1) model with the series selected in group G; call the coefficients ϕ_G and ψ_G;
4. verify the joint null hypothesis $\phi_G = \phi_i$ and $\psi_G = \psi_i$ for each remaining location i and, if at least one hypothesis is not rejected, select the series i which provides the maximum p-value, add it to the previous group and put $N^* = N^* - 1$;
5. if at least one hypothesis is not rejected, repeat steps 3. and 4.; if all the null hypotheses are rejected, the series selected until step 4. form a group of locations with the same coefficients.
6. repeat steps 2.–5. with the remaining series until no series remain.

It is important to underline some points.

There are two *counters* (N^* and G) to record the number of series remaining and the number of groups identified respectively. Notice that the number of groups G is identified by the algorithm and does not need to be fixed a priori. When, in step 2., we reject all the null hypotheses, we will obtain N^* groups of size 1.

If all the elements of the row i of matrix $\boldsymbol{W}$ are equal to zero, the previous hypotheses, involving location i, are relative only to the ϕ coefficient because the ψ coefficient is not identified and is a nuisance parameter.

In step 2. the p–value is adopted to establish the order in which the locations enter into the groups; this kind of approach does make sense in a clustering algorithm because, as shown by Maharaj (1999), the p–value is a measure of similarity and satisfies the properties of a semi–metric.

The test of the hypothesis in steps 2. and 4. can be performed by the Wald statistics. The validity of this approach in time series clustering has been verified by Otranto (2008 and 2010) with simulation experiments and it is based on the theory developed by Steece and Wood (1985). More in detail, let us suppose that, in step 4., we compare the parameters of a group of locations with the parameters of a location i; let us consider:

$$\boldsymbol{\theta} = (\phi_G, \psi_G, \phi_i, \psi_i)'; \qquad \boldsymbol{\Sigma} = \left[\begin{array}{cc} \boldsymbol{\Sigma}_G & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_i \end{array} \right] \tag{3.1}$$

where $\boldsymbol{\Sigma}_G$ is the covariance matrix of the parameters estimator of group G, $\boldsymbol{\Sigma}_i$ the covariance matrix of the parameters estimator of the location i and $\boldsymbol{0}$ a matrix with all the elements equal to 0. Each matrix has dimension (2×2). Moreover let:

$$\boldsymbol{C} = \left[\begin{array}{cccc} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{array} \right]$$

be a constant matrix; the Wald statistic is expressed by:

$$\Xi = (\boldsymbol{C}\boldsymbol{\theta})' \left(\boldsymbol{C}\boldsymbol{\Sigma}\boldsymbol{C}'\right)^{-1} (\boldsymbol{C}\boldsymbol{\theta}) \tag{3.2}$$

In step 2., the same statistic is adopted, but in (3.1) we have to substitute ϕ_G, ψ_G and Σ_G with ϕ_h, ψ_h and Σ_h, respectively.

4 Simulation Study

In order to verify the capability of the previous algorithm in detecting the correct groups of locations, we performed a simulation experiment, where we generate data from STAR(1,1) and FSTAR(1,1) models and then we apply the hierarchical algorithm. The data were generated from several combinations of time length T, number of locations N, number of groups G and spatial weight matrix $\boldsymbol{W}$. In particular we considered:

- three different time spans ($T = 100, 500, 1000$);
- two different spatial patterns, represented by regular lattices 3×3 ($N = 9$) and 5×5 ($N = 25$);
- three different number of groups ($G = 1, 2, 3$);
- three different spatial weight matrices, based on the classical contiguity criteria for regular lattices (rook, queen and bishop cases; see Anselin, 1988). These binary matrices are successively row standardized so that each row adds up to 1.

The data are generated from multivariate standard Normal distributions; we consider both the cases of uncorrelated disturbances $\boldsymbol{\varepsilon}_t$ and disturbances spatially correlated, transforming the generated $\boldsymbol{\varepsilon}_t$ by (2.2) with $\rho = 0.7$.

The coefficients used to generate data from the FSTAR model (2.5) are:
– when $G = 1$ (STAR model), $\phi_1 = 0.5$, $\psi_1 = 0.3$;
– when $G = 2$, $\phi_1 = 0.5$, $\psi_1 = 0.3$, $\phi_2 = 0.3$, $\psi_2 = 0.6$;
– when $G = 3$, $\phi_1 = 0.5$, $\psi_1 = 0.3$, $\phi_2 = 0.3$, $\psi_2 = 0.6$, $\phi_3 = 0.8$, $\psi_3 = 0.1$.
The number of replications is 1000 for each case.

The performance evaluation is conducted: 1) recording for each replication the number of groups detected; 2) verifying the similarity of the composition of the detected group with the true one. The second point is made because the number of groups could be correctly detected but the composition of the groups is not equal to the true one; moreover we are interested in evaluating the magnitude of the differences. For this purpose we adopt the adjusted Rand index (Rand, 1971; Hubert and Arabie, 1985):

$$r = \frac{\sum_{i=1}^{G}\sum_{j=1}^{G^*}\binom{\hat{n}_{ij}}{2} - [\sum_{i=1}^{G}\binom{n_i}{2}][\sum_{j=1}^{G^*}\binom{\hat{n}_j}{2}]/\binom{N}{2}}{[\sum_{i=1}^{G}\binom{n_i}{2} + \sum_{j=1}^{G^*}\binom{\hat{n}_j}{2}]/2 - [\sum_{i=1}^{G}\binom{n_i}{2}][\sum_{j=1}^{G^*}\binom{\hat{n}_j}{2}]/\binom{N}{2}} \tag{4.1}$$

where G and G^* represent the number of groups in the true and detected clustering respectively; n_i and $\hat{n}_j$ are the number of locations belonging to the group i of the true and group j of the detected clustering respectively, whereas $\hat{n}_{ij}$ is the number of locations belonging to the group i in the true pattern and assigned to the group j in the detected clustering. We can use r to evaluate the performance of the proposed method because $r \in [0, 1]$, assuming value 0 when the differences between the groups are at their maximum (worst performance) and 1 in the case of coincidence between the true and the detected clustering. We could be satisfied with high values of r, for example, more than 0.85.

In Tables 1–3 we show the results relative to the rook contiguity matrix;[2] each table refers to a different G and contains the results relative to the three different time spans, the two regular lattices and both the uncorrelated and correlated disturbances. When the data generating process is the STAR(1,1) model ($G = 1$; Table 1), the correct detection of the true model is very frequent in the 3×3 lattice independently on the time length: the Rand index (4.1) is equal to 1 in more than 83% of cases and this value increases with spatial autocorrelation of the disturbances. This percentage decreases sensitively in the 5×5 lattice, with a higher percentage of 2 groups detected, but the difference with respect to the true pattern is not relevant; in fact the percentage of cases with $r \geq 0.85$ is always around 96%. The procedure seems to fail when a FSTAR(1,1) model is the true data generating process and T is small. In Tables 2 ($G = 2$) and 3 ($G = 3$) we can notice the small number of cases with correct pattern detection, whereas the performance is good when T increases, with percentages of $r \geq 0.85$ near to 100%. It is necessary to point out also that, in the $T = 500$ and $T = 1000$ cases, when the number of groups is correctly detected (first two columns of the Tables), in general they are equal to the true generated patterns ($r = 1$).

In practice the algorithm does work for large T, whereas the correct pattern is difficult to be detected when T is small and $G > 1$. However, in practical terms, the important question is to understand if the FSTAR model (2.5) demonstrates a better fitting and better forecasting performance than the STAR model (2.1) also in the presence of small T, as well as a similar fitting and forecasting performance when compared to the USTAR model (2.4). This in principle would provide the best results. These aspects can be evaluated in a real space–time series, which is the purpose of next section.

5 Application

Let us consider the data set of Italian Crude Birth Rate (henceforth CBR) from January 2003 to October 2015 (154 monthly observations) and relative to the 20 Italian regions (Istat source: http://demo.istat.it). In order to obtain a stationary series, without losing observations,[3] we have subtracted a linear trend (of the form $a_i + b_i t$, $t = 1, \ldots, T$, $i = 1, \ldots, N$) to each time series. To verify the robustness of the results with respect to the $\boldsymbol{W}$ spatial matrix chosen, we have selected eight different weight specifications (for a review see Cliff and Ord, 1973, 1981; Getis and Ord, 1992; Cressie, 1993), distinguishing between binary and kernel spatial weight matrices:

- **Binary matrix**: the weight w_{ij} is equal to 1 if i and j are *neighbors*, 0 otherwise. The neighboring is defined in terms of boundaries or distance under a certain threshold.

[2] To save space we do not present the results for the queen and bishop case, which are available on request.

[3] This would happen using the difference operator.

1. *Boundary Matrix* (BM): it is the classical contiguity binary matrix based on the common boundary; in practice the weight is 1 when the two locations share a common boundary; by convention $w_{ii} = 0$ for $i = 1, \ldots, N$.
2. *Maxmin matrix* (Mm): the neighbors are detected by the Maxmin distance (Mucciardi and Bertuccelli, 2012). It is defined as

$$d_{Mm} = max(e_1, \ldots, e_N)$$

   where e_i ($i = 1, \ldots, N$) represents the minimum Euclidean distance of the generic location i and the other locations j ($i \neq j$). As a consequence all locations have at least one connection and the neighbors of location i are the locations with Euclidean distance lower than d_{Mm}.

- **Kernel matrix**: the weight w_{ij} is a continuous and monotonic decreasing function of the (Euclidean) distance d_{ij} (Fotheringham et. al., 2002). The choice of kernel functions is particularly appropriate because the bandwidth h provides a control about the circular area of influence of each observation i. We adopt Gaussian distance-decay-based functions as:

$$w_{ij} = exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{h}\right)^2\right]$$

  For our experiments we have selected the following kernel weight matrices:

3. $Kmin$: with $h = Min(d_{ij})$ (in our application it corresponds to 72 Kilometers) ;
4. $K10p$: with $h = 0.1Max(d_{ij})$ (106 KM);
5. $K20p$: with $h = 0.2Max(d_{ij})$ (213 KM);
6. KMm: with $h = d_{Mm}$ (379 KM);
7. $Kmea$: with $h = mean(d_{ij})$ (453 KM);
8. *Uniform matrix* (U): $h \to \infty$; this implies that $w_{ij} = 1$ for each $i \neq j$. It is the case in which each location is linked to each other and the distance between two locations is not relevant. It implies a model without a real spatial dependence, but a full interdependence among all locations.

In Figure 1 the behavior of the Maxmin distance and the first five kernel functions are illustrated.[4] The Maxmin binary criterion is a step function which assigns weight 1 to the locations with a Euclidean distance lower than the Maxmin distance; the Kernel weights decrease with the distance (in Kilometers) and the area of influence increases with h.

Each weight matrix is successively standardized by row.

We have estimated, separately for each weight matrix, the STAR(1,1), the USTAR(1,1) and the FSTAR(1,1) models. For the FSTAR model, we have applied the clustering algorithm with each weight matrix, obtaining different clustering. Using the Rand index (4.1)

[4]The U kernel function is not represented, but it would be a horizontal line in correspondence of the value 1.

for each pair of classifications, we obtain a synthetic representation of the different results derived from the different $\boldsymbol{W}$ matrices, as shown in Table 4. All the eight weight matrices provide different classifications, with maximum Rand index equal to 0.78, in correspondence of the pair KMm–$Kmea$; other similar classifications (with $r < 0.7$) are $Kmin$–U, $Kmin$–$K10p$ and BM–$Kmin$. On the other side there are several pairs with extremely low Rand index, with the minimum in correspondence of BM–Mm ($r = 0.18$). As expected, the BM matrix provides results more similar to kernel matrices with small h, whereas the U matrix provides results similar to kernel matrices with high h.

In Table 5 we show the estimates of USTAR and FSTAR models using the $Kmea$ weight matrix, which, in terms of Mean Squared Error (MSE), show the best results.[5] The horizontal lines separate the four groups identified by the hierarchical algorithm. Notice the non–significance of several USTAR coefficients; this is the typical problem of this model which makes it unfeasible in real cases. Also, the values of the ϕ and ψ coefficients estimated with the USTAR model are similar to the ϕ and ψ FSTAR coefficients of the corresponding group. The exceptions, such as Calabria in the first cluster, Umbria and Marche in the third cluster, are characterized by the highest standard errors, implying a strong inefficiency in the case of USTAR model.

The comparison of the models is performed in terms of MSE, separately for each region. We are essentially interested in investigating the presence of significative differences in the MSE of the three models; for this purpose we use the popular Diebold and Mariano (1995) test with the correction proposed by Harvey et al. (1997). In Figures 2 and 3 we show, in graphical terms, the results of this test (call it DM test), at a 5% significance level, for each weight matrix. Most of the regions show a similar behavior of the USTAR and FSTAR models, both significantly better than the STAR model (gray areas). There are a few cases with no prevalence of a model (white areas); in four cases this situation is relative only to one region (Calabria) and at least it covers four regions in the case of the $Kmin$ kernel matrix. Only in one case the USTAR model is significantly better than others (Friuli–Venezia Giulia with $\boldsymbol{W} = K10p$, denoted by a wireframe area).

These results indicate that the parsimonious FSTAR model, in general, shows an in–sample forecasting performance similar to the overparameterized USTAR model and better than the simple STAR model. This result is particularly interesting in view of the comments of Hansen (2010), who shows analitycally that the overparameterized models have a tendency to perform better than simpler models.

We also evaluate the performance of the three models in terms of out–of–sample performance. For this purpose we re-estimate the models on a reduced data set, excluding the last two years (from November 2013 to October 2015), re–applying the clustering algorithm to estimate the FSTAR model; then we perform the one–step ahead forecasts for the out–of–sample period (24×20 space–time observations). In this case, given the small number of forecasts for each region, we prefer to evaluate the full set of space–time forecasting squared errors. In Table 6 we show the out–of–sample MSE of the three models with the eight different weight matrices (left part of the table) and the corresponding p-values of the DM statistic (right part). In two cases (matrices Mm and $Kmea$) the FSTAR model has a lower MSE than USTAR and in one case ($K10p$) the two competitive

[5] The estimates of all the other models are available on request.

models show the same MSE. Observing the p–values, we notice that the out–of–sample forecasting performance of FSTAR and USTAR models is not significantly different at each significance level. The comparison with the STAR model favors USTAR in six cases and FSTAR in three cases both at a significance level of 5%; instead the comparison with the STAR model favors USTAR in six cases and FSTAR in seven cases at a significance level of 10%. We can conclude that the FSTAR model has a very similar behavior with respect to the USTAR model also in terms of out–of–sample performance.

6 Final Remarks

The use of a parsimonious model, as the STAR(1,1), to represent space–time series is a common practice in statistical modeling, also because the alternative VAR model (we called it USTAR model) causes inefficiency in the estimation and overparameterization. We have proposed a flexible model (FSTAR) with a reduced number of parameters, which allows regions with similar dynamics to have the same coefficients. The identification of the regions can be done through a hierarchical clustering algorithm, based on a Wald test which verifies the similarity of the coefficients of different locations. Consequently, the advantage of the FSTAR model is twofold: the model is a good compromise between the parsimony of the STAR model and the flexibility of the overparameterized USTAR model; the model building procedure identifies groups of locations with a similar space–time behavior and this result can be used for spatial aggregation or clustering analysis.

The simulation results show a good capability for identifying the true patterns for high time dimension, whereas this result is poor when the time length is small. Anyway, the in–sample and out–of–sample performance of the FSTAR model, in our practical application, seems to provide results similar to the overparameterized USTAR model and significantly better than the STAR model.

One of the crucial problems of the space–time models is the choice of the spatial weight matrix; in our application we have used eight exogenous weight matrices, which provide different clustering for the FSTAR model; anyway the results, in terms of in–sample and out–of–sample performance, seem sufficiently robust. Alternatively, in the case of kernel spatial matrices, we could estimate the bandwidth h with the other unknown coefficients, considering an exogenous $\boldsymbol{W}$ matrix (see Otranto et al., 2016).

Furthermore, the clustering and identification procedure can be extended to larger time and spatial lags, implying the consideration of a larger number of coefficients and a Wald test with several constraints to be jointly verified. Similar considerations could be made extending the methodology to STARMA models; in fact the clustering procedure is based on the Steece and Wood (1985) equivalence test, which was developed for general ARMA models. We leave these extensions to future research.

Acknowledgements

This research was supported by University of Messina Research&Mobility 2015 Project (project code RES_AND_MOB_2015_ DISTASO).

References

[1] Anselin, L. (1988): *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Dordrecht.

[2] Arbia, G., Bee, M. and Espa, G. (2011): Aggregation of regional economic time series with different spatial correlation structures. *Geographical Analysis* 43, 78–103.

[3] Besag, J. (1974): Spatial Interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B* 36, 192–236.

[4] Cliff, A. D. and Ord, J. K. (1973): *Spatial Autocorrelation.* Pion, London .

[5] Cliff, A.D. and Ord, J.K. (1975): Space–time modeling with an application to regional forecasting. *Transactions of the Institute of British Geographers* 64, 119–128.

[6] Cliff, A. D. and Ord, J. K. (1981): *Spatial Processes, Models and Applications.* Pion, London .

[7] Cressie, N. (1993): *Statistics for Spatial Data*. Wiley, New York.

[8] Diebold, F.X., and Mariano, R.S. (1995): Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13, 253–263.

[9] Fotheringham, AS., Brunsdon, C., and Charlton, ME. (2002): *Geographically Weighted Regression: the Analysis of Spatially Varying Relationship.* Wiley, New York.

[10] Getis, A. and Ord, J. K. (1981): The analysis of spatial association by use of distance statistics. *Geographical Analysis* 24, 189–206.

[11] Giacomini, R. and Granger, C.W.J. (2004): Aggregation of space–time processes. *Journal of Econometrics* 118, 7–26.

[12] Haining, R. (1990): *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, Cambridge.

[13] Hansen, P.R. (2010): A Winner's curse for econometric models: on the joint distribution of in–sample fit and out–of–sample fit and its implications for model selection. *Mimeo*, Stanford University.

[14] Harvey, D., Leybourne, S. and Newbold, P. (1997): Testing the equality of prediction mean squared errors. *International Journal of Forecasting* 13, 281–291.

[15] Hubert, L., Arabie, P. (1985): Comparing partitions. *Journal of Classification* 2, 193–218.

[16] LeSage, J. and Pace, R.K. (2009): *Introduction to Spatial Econometrics*. Chapman & Hall, Boca Raton.

[17] Liao, T. (2005). Clustering time series data: A survey. *Pattern Recognition* 38, 1857–1874.

[18] Lütkepohl, H. (1993): *Introduction to Multiple Time Series Analysis*. Springer Verlag, Berlin.

[19] Maharaj, E.A. (1999). Comparison and classification of stationary multivariate time series. *Pattern Recognition* 32, 1129–1138.

[20] Mucciardi, M. and Bertuccelli, P. (2012). The impact of the weight matrix on the local indicators of spatial association: an application to per–capita value added in Italy. *International Journal of Trade and Global Markets* 5, 133–141.

[21] Otranto, E. (2008): Clustering heteroskedastic time series by model–based procedures. *Computational Statistics and Data Analysis* 52, 4685–4698..

[22] Otranto, E. (2010): Identifying financial time series with similar dynamic conditional correlation. *Computational Statistics and Data Analysis* 54, 1–15.

[23] Otranto, E., Mucciardi, M. and Bertuccelli, P. (2016): Spatial effects in dynamic conditional correlations. *Journal of Applied Statistics* 43, 604–626.

[24] Pfeifer, P.E. and Deutsch, S.J. (1980): A Three–stage iterative procedure for space–time modeling. *Technometrics* 22, 35–47.

[25] Pinkse, J. and Slade, M.E. (1998): Contracting in space: an application of spatial statistics to discrete-choice models. *Journal of Econometrics* 85, 125–154.

[26] Rand, W.M. (1971): Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association* 66, 846–850.

[27] Steece, B. and Wood, S. (1985). A test for the equivalence of k ARMA models. *Empirical Economics* 10, 1–11.

[28] Stetzer, F. (1982). Specifying weights in spatial forecasting models: the results of some experiments. *Environment and Planning A* 14, 571–584.

Table 1: Simulation results for 3×3 and 5×5 lattices with the rook contiguity matrix and number of groups $n = 1$, different time length T, using uncorrelated ($\rho = 0$) and correlated ($\rho = 0.7$) disturbances.

			$T = 100; \rho = 0$				
Frequency			Rand index				
Lattice			3×3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	830	691	0	7	163	0	830
2	167	286			5×5 lattice		
3	3	23	0	3	39	267	691
			$T = 100; \rho = 0.7$				
Frequency			Rand index				
Lattice			3×3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	889	706	0	10	101	0	889
2	110	281			5×5 lattice		
3	1	13	0	3	39	252	706
			$T = 500; \rho = 0$				
Frequency			Rand index				
Lattice			3×3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	847	663	0	4	149	0	847
2	151	322			5×5 lattice		
3	2	15	0	2	42	293	663
			$T = 500; \rho = 0.7$				
Frequency			Rand index				
Lattice			3×3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	899	753	0	11	90	0	899
2	101	236			5×5 lattice		
3	0	11	0	4	34	209	753
			$T = 1000; \rho = 0$				
Frequency			Rand index				
Lattice			3×3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	863	693	0	7	130	0	863
2	134	291			5×5 lattice		
3	3	16	0	1	40	266	693
			$T = 1000; \rho = 0.7$				
Frequency			Rand index				
Lattice			3×3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	868	740	0	12	120	0	868
2	130	254			5×5 lattice		
3	2	6	0	0	40	220	740

Note: G^* indicates the number of groups detected by the algorithm described in Section 3. The Rand index r compares the true composition of the n groups and the one obtained from the algorithm described in Section 3. The STAR coefficients used to generate the data are $\phi_1 = 0.5$, $\psi_1 = 0.3$. The number of replications is 1000.

Table 2: Simulation results for 3×3 and 5×5 lattices with the rook contiguity matrix and number of groups $n = 2$, different time length T, using uncorrelated ($\rho = 0$) and correlated ($\rho = 0.7$) disturbances.

$T = 100; \rho = 0$							
Frequency			Rand index				
	Lattice		3 × 3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	212	71	224	594	122	4	56
2	718	734			5 × 5 lattice		
3	70	192	17	923	49	7	4
4	0	3					

$T = 100; \rho = 0.7$							
Frequency			Rand index				
	Lattice		3 × 3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	135	31	138	491	227	6	138
2	812	787			5 × 5 lattice		
3	52	177	12	726	169	69	24
4	1	5					

$T = 500; \rho = 0$							
Frequency			Rand index				
	Lattice		3 × 3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	0	0	1	11	100	49	839
2	859	658			5 × 5 lattice		
3	134	309	0	5	39	368	588
4	7	31					
5	0	2					

$T = 500; \rho = 0.7$							
Frequency			Rand index				
	Lattice		3 × 3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	0	0	0	11	76	47	866
2	868	716			5 × 5 lattice		
3	121	245	0	1	11	294	694
4	11	36					
5	0	3					

$T = 1000; \rho = 0$							
Frequency			Rand index				
	Lattice		3 × 3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	0	0	0	8	90	53	849
2	849	630			5 × 5 lattice		
3	143	325	0	0	3	367	630
4	7	43					
5	1	2					

$T = 1000; \rho = 0.7$							
Frequency			Rand index				
	Lattice		3 × 3 lattice				
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	0	0	0	12	58	51	879
2	879	688			5 × 5 lattice		
3	109	276	0	0	14	298	688
4	12	36					

Note: G^* indicates the number of groups detected by the algorithm described in Section 3. The Rand index r compares the true composition of the n groups and the one obtained from the algorithm described in Section 3. The STAR coefficients used to generate the data are $\phi_1 = 0.5$, $\phi_2 = 0.3$, $\psi_1 = 0.3$, $\psi_1 = 0.6$. The number of replications is 1000.

Table 3: Simulation results for 3×3 and 5×5 lattices with the rook contiguity matrix and number of groups $n = 3$, different time length T, using uncorrelated ($\rho = 0$) and correlated ($\rho = 0.7$) disturbances.

					$T = 100; \rho = 0$		
Frequency					Rand index		
	Lattice				3×3 lattice		
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	1	0	284	611	88	2	15
2	442	218			5×5 lattice		
3	524	642	167	790	37	4	2
4	31	127					
5	4	13					
					$T = 100; \rho = 0.7$		
Frequency					Rand index		
	Lattice				3×3 lattice		
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	0	0	230	548	166	1	55
2	252	121			5×5 lattice		
3	706	714	90	762	135	10	3
4	38	154					
5	3	11					
6	1	0					
					$T = 500; \rho = 0$		
Frequency					Rand index		
	Lattice				3×3 lattice		
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	0	0	0	6	24	128	842
2	0	0			5×5 lattice		
3	862	637	0	0	20	390	590
4	132	302					
5	5	58					
6	1	3					
					$T = 500; \rho = 0.7$		
Frequency					Rand index		
	Lattice				3×3 lattice		
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	0	0	0	1	14	73	912
2	0	0			5×5 lattice		
3	915	686	0	0	2	315	683
4	74	267					
5	11	44					
6	0	3					
					$T = 1000; \rho = 0$		
Frequency					Rand index		
	Lattice				3×3 lattice		
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	0	0	0	0	4	122	874
2	0	0			5×5 lattice		
3	874	624	0	0	5	371	624
4	122	318					
5	4	49					
6	0	9					
					$T = 1000; \rho = 0.7$		
Frequency					Rand index		
	Lattice				3×3 lattice		
G^*	3×3	5×5	$r < 0.5$	$0.5 \leq r < 0.7$	$0.7 \leq r < 0.85$	$0.85 \leq r < 1$	$r = 1$
1	0	0	0	0	1	66	933
2	0	0			5×5 lattice		
3	933	662	0	0	8	330	662
4	66	293					
5	1	42					
6	0	3					

Note: G^* indicates the number of groups detected by the algorithm described in Section 3. The Rand index r compares the true composition of the n groups and the one obtained from the algorithm described in Section 3. The STAR coefficients used to generate the data are $\phi_1 = 0.5$, $\phi_2 = 0.3$, $\phi_3 = 0.8$, $\psi_1 = 0.3$, $\psi_2 = 0.6$, $\psi_3 = 0.1$. The number of replications is 1000.

Table 4: CBR data set: Rand index for each pair of classification derived from the hierarchical algorithm with different spatial weight matrices.

	Mm	$Kmin$	$K10p$	$K20p$	KMm	$Kmea$	U
BM	0.182	0.716	0.583	0.346	0.325	0.302	0.270
Mm		0.174	0.296	0.448	0.480	0.531	0.412
$Kmin$			0.708	0.413	0.315	0.346	0.261
$K10p$				0.556	0.465	0.502	0.390
$K20p$					0.669	0.647	0.655
KMm						0.783	0.742
$Kmea$							0.670

Table 5: CBR data set: Estimation (standard errors in parentheses) of the coefficients of FSTAR and USTAR models, using the $Kmea$ weight matrix.

Region	**ϕ FSTAR**		**ϕ USTAR**		**ψ FSTAR**		**ψ USTAR**	
Liguria	0.747	(0.119)	0.757	(0.220)	-0.562	(0.154)	-0.626	(0.355)
Molise			0.730	(0.267)			-0.592	(0.369)
Basilicata			0.668	(0.376)			-0.495	(0.415)
Calabria			1.053	(2.056)			-0.758	(2.094)
Sardinia			0.725	(0.315)			-0.583	(0.401)
Trentino Alto-Adige	1.145	(0.110)	1.127	(0.189)	-0.997	(0.242)	-0.992	(0.513)
Lazio			1.263	(0.412)			-1.115	(0.672)
Campania			1.115	(0.211)			-0.986	(0.522)
Sicily			1.217	(0.352)			-1.076	(0.625)
Piedmont	0.353	(0.279)	0.133	(0.890)	-0.187	(0.255)	-0.006	(0.784)
Friuli-Venezia Giulia			0.498	(0.464)			-0.364	(0.431)
Tuscany			0.203	(0.883)			-0.075	(0.794)
Umbria			-0.542	(1.793)			0.673	(1.660)
Marche			-0.718	(2.336)			0.895	(2.218)
Abruzzo			0.315	(0.768)			-0.179	(0.697)
Aosta Valley	1.430	(0.232)	1.412	(0.512)	-1.271	(0.330)	-1.283	(0.781)
Lombardy			1.337	(0.423)			-1.202	(0.702)
Veneto			1.495	(0.587)			-1.364	(0.853)
Emilia-Romagna			1.674	(1.029)			-1.480	(1.223)
Apulia			1.740	(0.992)			-1.597	(1.212)

Note: The horizontal lines separate the groups identified by the procedure described in section 3. The parameters estimated with the STAR(1,1) model are $\phi = 0.957\ (0.070)$ and $\psi = -0.670\ (0.102)$.

Table 6: Out–of–sample forecasting of the CBR of the Italian Regions: MSE (multiplied by 100) of STAR, USTAR and FSTAR models, and p-value of the DM statistic comparing the means of the squared errors of each pair of models.

	MSE			p-value DM statistic		
W	STAR	USTAR	FSTAR	STAR vs USTAR	STAR vs FSTAR	USTAR vs FSTAR
BM	0.357	0.303	0.311	0.008	0.018	0.356
Mm	0.369	0.321	0.300	0.015	0.000	0.162
$Kmin$	0.360	0.303	0.328	0.000	0.071	0.126
$K10p$	0.363	0.338	0.338	0.164	0.097	0.488
$K20p$	0.358	0.328	0.348	0.111	0.348	0.156
KMm	0.350	0.312	0.333	0.007	0.195	0.154
$Kmea$	0.348	0.306	0.283	0.000	0.000	0.118
U	0.346	0.304	0.322	0.000	0.084	0.185

Note: The estimation is performed on the period January 2003 – October 2013; the out–of–sample span is November 2013 – October 2015. The results are referred to the 24×20 space–time forecasts.

Figure 1: Weights of the spatial matrix derived from the Maxmin criterion and five Kernel functions.

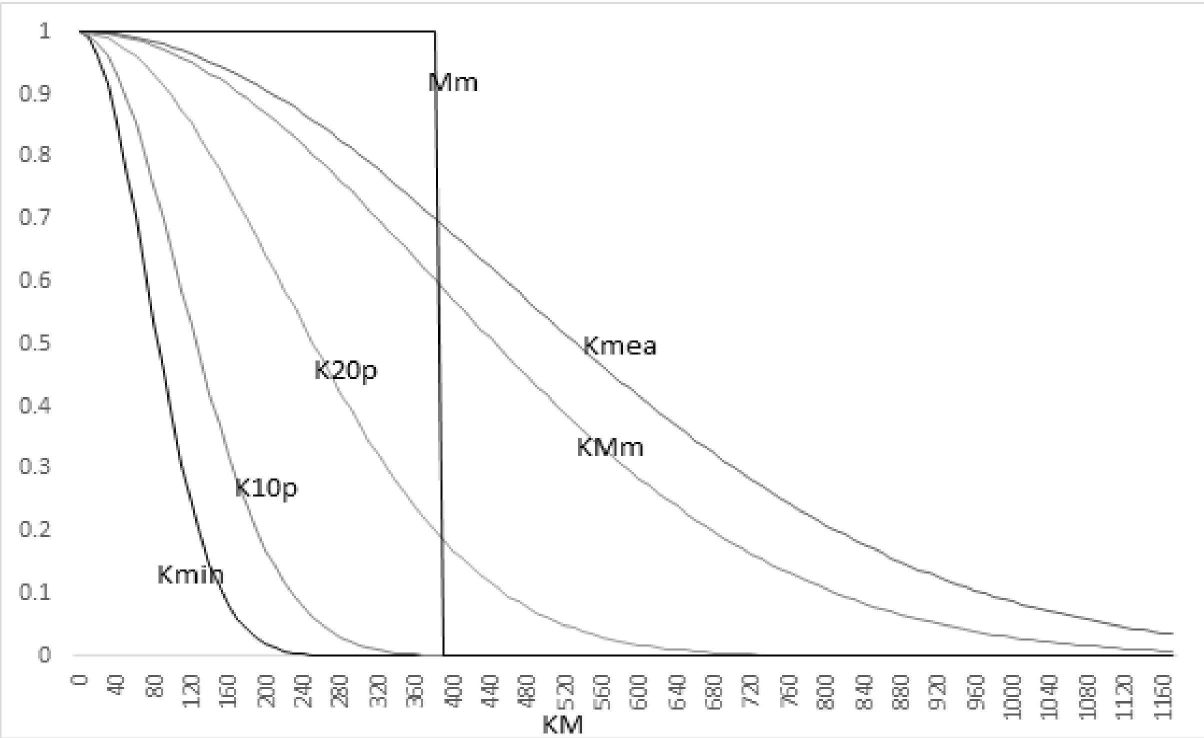
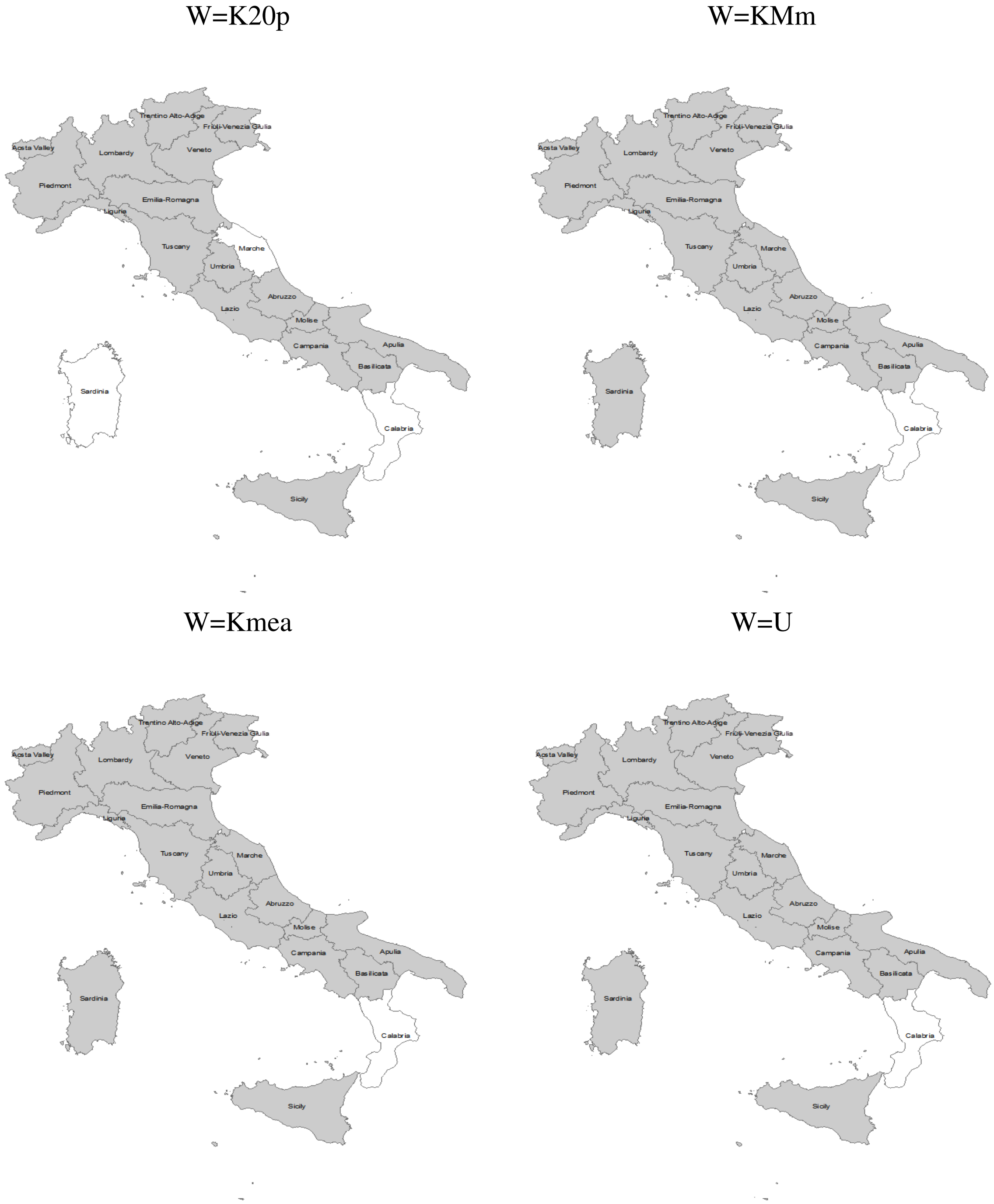


Figure 2: Results of the DM test at 5% nominal size for the CBR of the Italian regions, comparing alternative STAR models with different weight matrix W (BM, Mm, $Kmin$, $K10p$.

Note:Gray colors indicate that, for the corresponding region, the USTAR and FSTAR models have the same in–sample forecasting performance, which is significantly better than the STAR model; white areas indicate that, for the corresponding regions, the three models have the same in–sample forecasting performance; the *wireframe* areas indicate that, for the corresponding regions, model USTAR has a significantly better in–sample forecasting performance than STAR and FSTAR models.

Figure 3: Results of the DM test at 5% nominal size for the CBR of the Italian regions, comparing alternative STAR models with different weight matrix W ($K20p$, KMm, $Kmea, U$.



Note:Gray colors indicate that, for the corresponding region, the USTAR and FSTAR models have the same in–sample forecasting performance, which is significantly better than the STAR model; white areas indicate that, for the corresponding regions, the three models have the same in–sample forecasting performance.

Ultimi Contributi di Ricerca CRENoS

I Paper sono disponibili in: http://www.crenos.it

- **16/07** *Leonardo Becchetti, Vittorio Pelligra, Francesco Salustri,* "Testing for Heterogeneity of Preferences in Randomized Experiments: A Satisfaction-Based Approach Applied to Multiplayer Prisoners' Dilemmas"
- **16/06** *Anna Bottasso, Maurizio Conti, Giovanni Sulis,* "Firm Dynamics and Employment Protection: Evidence from Sectoral Data"
- **16/05** *Oliviero A. Carboni, Giuseppe Medda,* "R&D, Export, and Investment Decision"
- **16/04** *Leonardo Becchetti, Germana Corrado, Vittorio Pelligra, Fiammetta Rossetti,* "Happiness and Preferences in a Legality Social Dilemma Comparing the Direct and Indirect Approach"
- **16/03** *Leonardo Becchetti, Vittorio Pelligra, Francesco Salustri, Alejandra Vásquez* "Gender differences in Socially Responsible Consumption. An Experimental Investigation"
- **16/02** *Leonardo Becchetti, Vittorio Pelligra, Alejandra Vásquez,* "Are Women *Naturaliter* More Cooperative? An Experimental Investigation of the Vote-with-the-Wallet Game"
- **16/01** *Leonardo Becchetti, Vittorio Pelligra, Fiammetta Rossetti,* "The Corporate Legality Game A Lab Experiment on The Impact of Policies, Frames and Information"
- **15/22** *Fabio Cerina, Luisanna Cocco, Katiuscia Mannaro, Michele Marchesi, Francesco Pigliaru,* "Insularity and the development of a local network: a simulation model applied to the Italian railway system"
- **15/21** *Fabio Cerina,* "Is insularity a locational disadvantage? Insights from the New Economic Geography"
- **15/20** *Giorgio Garau, Giovanni Mandras,* "Economy-wide rebound effects from an increase in efficiency in the use of energy: the Italian case"
- **15/19** *Bianca Biagi, Maria Gabriela Ladu, Marta Meleddu, Vicente Royuela,* "Tourism and quality of life: a capability approach"
- **15/18** *Gianfranco Atzeni, Francesco Chessa, Luca G. Deidda, Malika Hamadi, Stefano Usai,* "Access to microcredit and borrowers' behavior: Evidence from Sardinia"
- **15/17** *Guerino Ardizzi, Federico Crudu, Carmelo Petraglia,* "The Impact of Electronic Payments on Bank Cost Efficiency: Nonparametric Evidence"
- **15/16** *Raffaele Brancati, Emanuela Marrocu, Manuel Romagnoli, Stefano Usai,* "Innovation activities and learning processes in the crisis. Evidence from Italian export in manufacturing and services"
- **15/15** *Oliviero A. Carboni, Giuseppe Medda,* "R&D Spending and Investment Decision: Evidence from European Firms"
- **15/14** *Leonardo Becchetti, Vittorio Pelligra, Serena F. Taurino,* "Other-Regarding Preferences and Reciprocity: Insights from Experimental Findings and Satisfaction Data"
- **15/13** *Luisanna Cocco, Manuela Deidda, Michele Marchesi, Francesco Pigliaru,* "Can islands profit from economies of density? An application to the retail sector?"

Finito di stampare nel mese di Dicembre 2016
Presso Centro Stampa dell'Università degli Studi di Cagliari
Via Università 40
09125 Cagliari

www.crenos.it

ISBN 978-88-9386-001-7