Spatial Panel Data Forecasting over Different Horizons, Cross-Sectional and Temporal Dimensions

Matías Mayor[1] and Roberto Patuelli[2]

[1] Department of Applied Economics, University of Oviedo, Spain, mmayorf@uniovi.es

[2] Department of Economics, University of Bologna, Italy; The Rimini Centre for Economic Analysis (RCEA), Italy, roberto.patuelli@unibo.it

Abstract

Empirical assessments of the forecasting power of spatial panel data econometric models are still scarcely available. Moreover, several methodological contributions rely on simulated data to showcase the potential of proposed methods. While simulations may be useful to evaluate the properties of a single estimator, the empirical set-ups of simulation studies are often based on strong assumptions regarding the shape and regularity of the statistical distribution of the variables involved. It is then valuable to have, next to simulation studies, empirical assessments of competing econometric models based on real data. In this paper, we evaluate competing spatial (dynamic) panel methods, selecting a number of data sets characterized by a range of different cross-sectional and temporal dimensions, as well as different levels of spatial autocorrelation. We carry out our empirical exercise on regional unemployment data for France, Spain and Switzerland. Additionally, we test different forecasting horizons, in order to investigate the speed of deterioration of forecasting quality. We compare two classes of methods: spatial vector autoregressive (SVAR) models and dynamic panel models making use of eigenvector spatial filtering (SF). We find that, as it could be expected, the unbalance between the temporal and cross-sectional dimension ($T >> n$) does play in favour of the SVAR model. On the other hand, the advantage of the SVAR model over the SF model appears to diminish as the forecasting horizon widens, eventually leading the SF model to being preferred for more distant forecasts.

1. Introduction

Labour market figures represent key information from a macro- and micro-economic point of view. In particular, being able to forecast such values – in the short- or long-run, depending on the problem at hand – is a valuable asset, when evaluating past and potential labour market policies and more. Barca et al. (2012) recently stressed the need for *local* (versus *place-neutral*) socio-economic policies, in the view of an improvement in their localized effectiveness and efficiency. The design of such local policies, and in particular active labour market policies, demands accurate social and economic information about regional endowments and all idiosyncratic factors that may affect their success. In order to allow policymakers to allocate public expenditure efficiently between regions, labour market forecasts at the regional level are a necessary complement to forecasts at the national level. This idea reinforces the need for econometric techniques allowing to obtain accurate local predictions, where, in particular, the information on each spatial unit is considered together with the one on its

neighbouring units, which are most likely to influence a spatial unit's socioeconomic system with their own trends, shocks and policies because of spatial proximity (e.g., through interregional commuting).

The existence – and persistence – of regional unemployment differences (Blanchard and Katz 1992; Partridge and Rickman 1995; Patacchini and Zenou 2007) has a double dimension: temporal and spatial. Then, the analyst faces the question of how this phenomenon may be incorporated in the forecasting strategy. The final aim is to obtain regional/local predictions taking into account, in addition to national and international trends, the fact that neighbouring regions may be structurally different and that, at the same time, it is necessary to incorporate in the models the existence of spatial economic linkages between regions (Longhi and Nijkamp 2007).

Empirical assessments of the forecasting power of spatial econometric models are still scarcely available. Moreover, several methodological contributions rely on simulated data to showcase the potential of proposed methods. While these are useful to evaluate the properties of the estimators, the empirical set-ups of simulation studies are often based on strong assumptions regarding the shape and regularity of the statistical distribution of the variables involved. It is then worth having, next to simulation studies, an assessment of the forecasting models on the basis of real data.

In this paper, we evaluate two competing spatial methods on a number of data sets characterized by a range of different cross-sectional and temporal dimensions. Additionally, the different levels of spatial autocorrelation of the data are explicitly considered. We carry out our empirical exercise on regional unemployment data for France, Spain and Switzerland. Additionally, we test different forecasting horizons, in order to investigate the speed of deterioration of the forecasting quality. We compare two classes of methods: spatial vector-autoregressive (SVAR) models and dynamic panel models augmented by eigenvector spatial filtering (SF). The two models chosen belong to two separate traditions: VAR models represent the mainstream (time-series) forecasting tradition, while SF-enhanced dynamic panel models attempt to merge the panel data modelling tradition to the spatial statistics one, within a semi-parametric framework. Both allow to inspect the spatial heterogeneity in the persistence of the regional unemployment rates (and therefore in the shock absorption speed), and to avoid imposing a unique coefficient for all regions (Longhi and Nijkamp 2007).

In (anonymized 2012), we analysed the short-run (one-period-ahead) forecasting performance of the above competing spatial models. The SVAR models showed somehow superior performance when the time dimension dominated, consistently with the time-series framework of VAR models. Moving instead to moderate cross-sectional and temporal dimensions, no clear difference could be drawn between the SVAR and SF estimation frameworks. We now focus on investigating how the performance of the two methods evolves over longer forecasting horizons. Our results point, consistently with previous findings, to median and average error levels that are tied to the data structure (cross-sectional and temporal dimensions). Sign tests comparing the two methods' forecasting errors appear to prefer the SVAR, but for longer forecasting horizons, the SF approach becomes competitive. The obtained forecasting errors of the SF models show stronger residual spatial autocorrelation (though still limited), while SVAR forecasting errors often exhibit negative spatial autocorrelation, suggesting that further fine-tuning is necessary in order to obtain the best possible performance from the two methods.

The paper is structured as follows. In Section 2, we provide a brief methodological description of the SVAR and SF, their benefits and limitations. In Section 3, we illustrate the data sets used and their characteristics. In Section 4, we present and discuss our empirical results. Finally, in Section 5 we draw some concluding remarks and outline future expansions of our empirical exercise.

2. Methodology: Spatial Vector-Autoregressive Models and Spatial Filtering-Aumented Dynamic Panel Data Models

The importance of the inclusion of spatial dependence and heterogeneity for forecasting purposes has been analysed in several papers. Giacomini and Granger (2004, p. 7) state that 'ignoring spatial autocorrelation, even when it is weak, leads to highly inaccurate forecasts'. Similarly, Hernández-Murillo and Owyang (2006) find that obtaining forecasts for disaggregated data using a space-time autoregressive model may lead to reductions in the out-of-sample mean squared error (MSE). Different methodological approaches have been proposed in the literature. Baltagi and Li (2004, 2006), Longhi and Nijkamp (2007), Fingleton (2009) and Baltagi et al. (2012) use static spatial panel data models. Kholodilin et al. (2008) and Kholodilin and Mense (2012), instead, use dynamic spatial panel data models, and find that accounting for spatial effects improves forecasting performance, in particular when the forecasting horizon is longer. Schanne et al. (2010) reach a similar conclusion, comparing a spatial global VAR (GVAR) model with time series methods. Angulo and Trívez (2010) find a substantial equivalence between a fixed effects panel spatial lag model and a series of non-spatial ARIMA models. On the other hand, Ohtsuka and Kakamu (2013) find that a VAR model outperforms a spatial autoregressive ARMA (SAR-ARMA) model.

In this paper, we aim to exploit the strong heterogeneity, across countries, in the size (e.g., in terms of population or area) of regions at the same official level of aggregation to investigate the forecasting performance of competing spatial econometric methods. We focus on two particular ones: an SVAR model proposed by Beenstock and Felsenstein (2007), and a dynamic heterogeneous-coefficients panel data model based on eigenvector SF (Griffith 2000; 2003, anonymized 2012). We deliberately select two models belonging to two separate traditions: VAR models come from the time-series forecasting tradition, and are widely used, in macroeconomics, to study adjustment processes; the SF-augmented dynamic panel data model connects the panel data modelling tradition [e.g., least squares dummy variables (LSDV) models] to the spatial statistics one, within a semi-parametric framework.

VAR models (Sims 1980) can be written as a set of symmetric equations in which each (dependent) variable is described by a set of its own lags and the lags of other variables in the system. VAR models are vastly employed to study, within a flexible estimation framework, the linkages between several variables. Restrictions are not imposed by theory but rather by statistical tools, as in VAR models the number of parameters to be estimated grows more than proportionally with the problem dimensionality. However, the flexibility of VAR models is only certain in the temporal dimension.[1] A standard VAR model does not incorporate spatial spillovers: for example, a shock in one region only influences the economic behaviour of the region itself.

---

[1] Structural VAR models are introduced to incorporate the necessary restrictions.

There are several proposals on how to account for spatial dependence (or, more generally, cross-sectional dependence) between regions in a VAR model. The most relevant obstacle in this regard is given by the fact that the number of parameters to be collected to model such relationships increases quadratically with the number of spatial units. In this regard, information on spatial proximity is used to limit the number of estimated parameters. Pan and LeSage (1995) propose to use spatial contiguity information as an alternative prior in a Bayesian VAR model. Similarly, Di Giacinto (2003) sets up parameter constraints in a structural VAR model on the basis of the neighbours structure, allowing to identify and estimate an SVAR model. Schanne et al. (2010) start from the GVAR model of Pesaran et al. (2004) and use geographical information to model spatial connections between regions. In particular, the GVAR model allows to include a temporal dimension within the spatial dependence process. In this regard, some authors consider only contemporaneous spatial processes (Longhi and Nijkamp 2007; Kholodilin et al. 2008), while others specify only temporally lagged spatial dependence (Hernández-Murillo and Owyang 2006).

All the methodological approaches mentioned above use spatial weights matrices, which are positive, (usually) non-stochastic matrices whose elements show the intensity of interdependence between pairs of spatial units. In our specification, we follow the SVAR approach proposed by Beenstock and Felsenstein (2007), who bring together mainstream VAR methods and spatial panel data techniques. Beenstock and Felsenstein allow for both contemporaneous and serially lagged spatially correlated variables. Since their SVAR model is highly nonlinear, because of contemporaneous spatial dependence, they restrict the coefficient of the contemporaneous spatial lags to zero, linearizing the model.

Let us consider n regions or municipalities ($i=1,\ldots,n$) where the values of a set of random variables are observed over time $t=1,\ldots,T$. In general terms, if we consider p temporal lags and s cross-regressive spatial lags, we must manage n equations (one for each region) like the following:

$$y_{i,t} = c_i + \sum_{p=1}^{P} \beta_{i,p} y_{i,t-p} + \sum_{s=1}^{S} \delta_{i,s} \sum_{j=1}^{n} w_{s,i,j} y_{j,t} + \sum_{s=1}^{S} \sum_{p=1}^{P} \gamma_{i,s,p} \sum_{j=1}^{n} w_{s,i,j} y_{j,t-p} + \varepsilon_{i,t}, \qquad (1)$$

where $w_{s,i,j}$ is the cell (i, j) of spatial weight matrix $\mathbf{W_s}$ for the sth spatial lag. The novelty of this model is the inclusion of the spatial cross-regressive lags, which are obtained by premultiplying each serially lagged variable by the spatial weight matrix. In the estimation stage, for each spatial unit a unique value of the spatial lag variable is estimated, and each observation has its own set of neighbouring units. Since $\sum_{j=1}^{n} w_{s,i,j} y_{j,t}$ and $\varepsilon_{i,t}$ are not independent, Equation (1) is estimated by means of seemingly unrelated regressions (SUR), where the coefficients of the contemporary spatial lags are estimated using, as instruments, the (spatially weighted) predicted values of the dependent variable. This is the common solution to the endogeneity issue caused by the spatial lag of the dependent variable in a simple spatial lag model.

If p = 1, each equation includes a constant, the serially lagged variable, and the contemporaneous and serially lagged spatial lags:

$$y_{i,t} = c_i + \beta_{i,1} y_{i,t-1} + \delta_{i,1} \sum_{j=1}^{n} w_{1,i,j} y_{j,t} + \gamma_{i,1,1} \sum_{j=1}^{n} w_{1,i,j} y_{j,t-1} + \varepsilon_{i,t}. \qquad (2)$$

In Equation (2), $\sum_{j=1}^{n} w_{1,i,j} y_{j,t}$ collects the contemporaneous spatial lags, and $\sum_{j=1}^{n} w_{1,i,j} y_{j,t-1}$ the serially lagged spatial lags.

In this paper, we compare the above method with the one recently proposed by (anonymized 2012), based on a heterogeneous-coefficients dynamic panel data model augmented by means of SF. This latter approach allows to account for spatial heterogeneity and/or autocorrelation both in the levels and in the regression coefficients, including the serial autoregressive term. Eigenvector SF (Griffith 2000, 2003) is a nonparametric solution to the problem of spatial autocorrelation in regression models. It relies on the computational formula of Moran's I (MI, Moran 1948), the most common indicator of spatial autocorrelation. After pre- and post-multiplying a spatial weights matrix $\mathbf{W}$ by a projection matrix, we obtain:

$$\mathbf{C} = (\mathbf{I}_n - \mathbf{1}\mathbf{1}^\mathrm{T} / n) \mathbf{W} (\mathbf{I}_n - \mathbf{1}\mathbf{1}^\mathrm{T} / n), \qquad (3)$$

where $\mathbf{1}$ is an n x 1 vector containing 1's. Matrix $\mathbf{C}$ can be used to obtain, given a variable X, the numerator of MI, and its extreme eigenvalues are approximately the extreme values of MI (Griffith 2000). Because of the mathematical relation between $\mathbf{C}$ and MI, the eigenvectors extracted from $\mathbf{C}$ represent all mutually exclusive (orthogonal and independent) spatial patterns implied by $\mathbf{W}$. The eigenvectors $E_1...E_n$ of $\mathbf{C}$ are extracted in decreasing order of spatial autocorrelation (according to MI). Therefore, E_1 has the largest MI achievable [$I(E_1)$], given $\mathbf{W}$. All subsequent eigenvectors maximize MI while being orthogonal to the eigenvectors previously extracted.

When employed in a regression model as additional explanatory variables, the above eigenvectors may account, among other things, for unobserved heterogeneity, redundant information, and spatial spillover effects, rendering regression residuals spatially uncorrelated (at least in a cross-sectional framework). A stepwise regression approach may be used to select which eigenvectors are actually significant in a specific modelling exercise. Because the number of eigenvectors increases with the cross-sectional dimension, it is customary to start the stepwise model selection from a subset of so-called 'candidate' (or 'dominant') eigenvectors. This subset is usually defined according to a minimum threshold of 0.25 for the ratio $I(E_j)/I(E_1)$ (for details, see Griffith 2003). The linear combination of the k eigenvectors selected in the stepwise procedure and their estimated coefficients is called a 'spatial filter'. Because of its basis in eigenvector decomposition, SF is close to principal components analysis, and when applied in a panel data modelling framework, the estimated spatial filter may be seen as a substitute for the individual fixed effects (anonymized 2012). In this regard, SF may be thought to be similar to the method recently brought forward by Bai (2013), who proposes a factor-analytical approach to the estimation of individual fixed effects.

Additionally, SF may be employed to inspect the spatial heterogeneity of regression coefficients (other than the intercept) as well, similarly to what is done in

geographically weighted regression (GWR, Fotheringham et al. 2002).[2] (anonymized 2012) show that, in dynamic panel data modelling, a heterogeneous-coefficients model can be effectively and efficiently approximated by constructing an SF representation of the vector of serial autoregressive coefficients. This is done by interacting each candidate eigenvector (piled T times to allow for the temporal dimension) with the serially lagged dependent variable, constructing a set of new variables that represent its decomposition in orthogonal spatial components. The regression coefficients estimated for the new variables will indicate the relevance of each spatial pattern in determining the spatial heterogeneity in the adjustment process (shock absorption) speed. The same process can be applied to any other explanatory variable, as in GWR. Additionally, simulations (anonymized 2013) show that the resulting estimator provides, for most combinations of n and T, smaller bias in the estimation of the autoregressive coefficient than the two standard methods in dynamic panel data modelling, the bias-corrected LSDV estimator (Bun and Carree 2005) and, for the homogeneous coefficients case, the system GMM estimator (Blundell and Bond 1998).

When spatial filters are simultaneously applied at the serial correlation and the intercept level, the following model is obtained:

$$y_{i,t} = c + \beta y_{i,t-1} + \sum_{m=1}^{k} \beta_m E_{i,m} y_{i,t-1} + \sum_{m'=1}^{k'} \beta_{m'} E_{i,m'} + \varepsilon_{i,t}, \qquad (4)$$

where k and k' are the number of eigenvectors selected at the lagged term and intercept level, respectively. A (standard) intercept c and an average serial correlation coefficient β can still be estimated, and the two spatial filters show the regional deviations from these aggregate measures due to spatial relationships.

A further advantage of the SF approach is that, unlike the SVAR, it can be used for any combination of cross-sectional (n) and temporal (T) dimensions, including the small-n small-T and big-n small-T cases (see, e.g., Giacomini and Granger 2004). For the latter in particular (e.g., the case for German NUTS-3 unemployment data), the SVAR model cannot be estimated. Therefore, these cases are, for the moment, excluded from our comparative analysis.

3. Data Description and Forecasting Strategy

The aim of this paper is to evaluate the forecasting performance of the spatial panel data methods described above using real-world data sets. We use data on official unemployment rates in France, Spain and Switzerland at the NUTS-3 level of geographical aggregation, and analyse in comparative terms the speed of deterioration of forecasting quality as the forecast horizon increases.

Previous studies have analysed '[h]ow far ahead into the future … forecasts have value, and how the information content of forecasts changes over forecast horizons' (Isiklar and Lahiri 2007, p. 167), but focusing only on the time domain. Using GDP forecasts, Isiklar and Lahiri find that forecasting error variability increases as the forecast horizon

---

[2] In this regard, Griffith (2008) shows that the SF-based approach to GWR provides superior statistical properties (e.g., with regard to multicollinearity) than the actual GWR.

increases, while forecast variability decreases. As a consequence, uncertainty (over a wide forecasting horizon) is associated with a great variability in the forecasting errors but less variability in the forecasting values. This is a further dimension along which it is worth comparing the SVAR and SF methods.

3.1. Data Description

In our empirical exercise, we initially consider four real-world data sets, characterized by a range of different cross-sectional and temporal dimensions, for France, Germany, Spain and Switzerland. The data for France, Spain and Switzerland have satisfactory but different temporal (T) and spatial dimensions (n), whereas German data have a big n (439) but a small T (36), so that the SVAR approach becomes unfeasible (the total number of parameters to be estimated in the model largely exceeds the total number of observations available[3]). Consequently, we decide to focus our empirical application on the three aforementioned countries.

Besides differences in the temporal and spatial dimensions, a further interesting aspect is that the size of the NUTS-3 spatial units for our reference countries differs widely. Then, it may be interesting to examine whether these differences affect the forecasting performance and in which direction. NUTS-3 spatial units correspond, for Spain, France and Switzerland, to provinces, departments and cantons, respectively. All of them are official delimitations, although with different powers and levels of autonomy. French departments have elected councils and administer a number of social welfare duties. Spanish provinces, instead, are mostly used for geographical reference and for delineating electoral districts. On the other hand, Swiss cantons are much more independent, each having its own constitution and government (as they used to be independent microstates). The average area of French departments is 7,030 km^2 (σ = 8,156.01 km^2), the one of Spanish provinces is 10,499 km^2 (σ = 4,699.77 km^2), while the same for Swiss cantons is 1,582 km^2 (σ = 1,822.35 km^2). Therefore, the average size of the Swiss cantons is clearly the lowest, and the highest level of variability is seen for French departments. Table 1 provides summary information on the spatial and temporal dimensions of our data sets.

Table 1. Data sets analysed in terms of their temporal and spatial dimensions

Country	n vs T dimension	Average size (std. deviation)
Spain (n = 47, T = 132)	$T \gg n$	10,499 km^2 (σ = 4,699.77 km^2)
France (n = 96, T = 120)	$T > n$	7,030 km^2 (σ = 8,156.01 km^2)
Switzerland (n = 25, T = 384)	$T \gg n$	1,582 km^2 (σ = 1,822.35 km^2)

Unemployment data for Spain are collected through the Spanish Labour Force Survey (EPA). The data consist of quarterly unemployment rates and cover the period 1976–2008. Most studies about Spanish labour markets assert that one of their main features is

---

[3] Canova and Ciccarelli (2013) analyse the technical options to confront the curse of dimensionality issue in the case of large panel VAR models.

the unemployment persistence from an aggregate viewpoint, but it should be noted that the persistence in the inequality of unemployment rates across provinces has been highlighted as well (Blanchard and Jimeno 1995; Jimeno and Bentolila 1998). For example, in 2006 some provinces experienced unemployment rates above 14 per cent (Cádiz, Badajoz, Huelva and Cordoba), whereas others (Teruel, Soria, Navarra and Guipuzcua) had rates lower than 6 per cent. These relevant differences are quite stable over the entire period studied.

Unemployment data for France are provided by the National Institute of Statistics and Economic Studies (INSEE) and consist of quarterly unemployment rates covering the period 1982–2011. The average unemployment rate for this period is 8.86 per cent and its temporal evolution is quite stable, showing its highest rate of 10.8 per cent in 1997 and its lowest (6.8 per cent) at the beginning of the period considered.

Switzerland is a non-EU country, and its labour market can be considered to be quite different from the EU average, although some constraints have been relaxed with regards to employment and migration regulations (Switzerland has recently abolished immigration quotas, and initiated participation in the Schengen agreement). Nonetheless, the Swiss labour market is still strictly regulated, and migration is controlled through working permits. Unemployment data for Switzerland are given through the Unemployment Statistics of the Swiss Federal Statistical Office. Our data set consists of monthly unemployment rates between 1975 and 2008, collected for the 26 cantons of Switzerland. Unemployment figures in Switzerland are much lower than in Spain and France. From 1995 to 2010, Switzerland's unemployment rate averaged 3.38 per cent, reaching an historical high of 5.40 per cent in 1997, and recording its lowest value (1.60 per cent) in 2000.

The difference between Switzerland's historical high and unemployment rates in Spain is striking. The temporal evolution of unemployment rates is also quite different, and Spanish data show a higher level of volatility. The French data represent an intermediate case between Spain and Switzerland, both in terms of average unemployment rates and in terms of cross-sectional and temporal dimensions, as shown in Figure 1. The plots composing the figure cover mostly overlapping periods, and employ, for each region (i.e., each row in a graph), a specific quantile-based colour scale. The bottom time series plot in each graph shows the evolution of the median regional unemployment rate. As it can be seen, most Spanish and French regions experienced two highs in unemployment, around 1985 and 1995, followed by a marked improvement and, ultimately, by the first signs of unemployment increase led by the 2008 financial crisis. Swiss regions, instead, experienced a marked unemployment increase between 1990 and 1995, which lasted until about 2000. It is worth noting that, while Swiss regions all follow the aggregate trend closely (the lighter and darker parts of the plot are homogeneous), the patterns for Spain and France are more irregular, suggesting possible heterogeneity in cyclical sensitivity, which, if not random, could be reflected in spatial clustering of serial correlation coefficients.

All three countries show spatially autocorrelated unemployment rates, as shown in Figure 2 in terms of MI. In Spain, the 1980s recorded a more homogeneous scenario, while recent trends appear to lead – as in the very beginning of our observation period – to much stronger clustering of unemployment rates, which identifies the highest values of MI over the three data sets. A similar pattern is shown for France, although the overall range of variation of MI is limited. In Switzerland, it appears that spatial

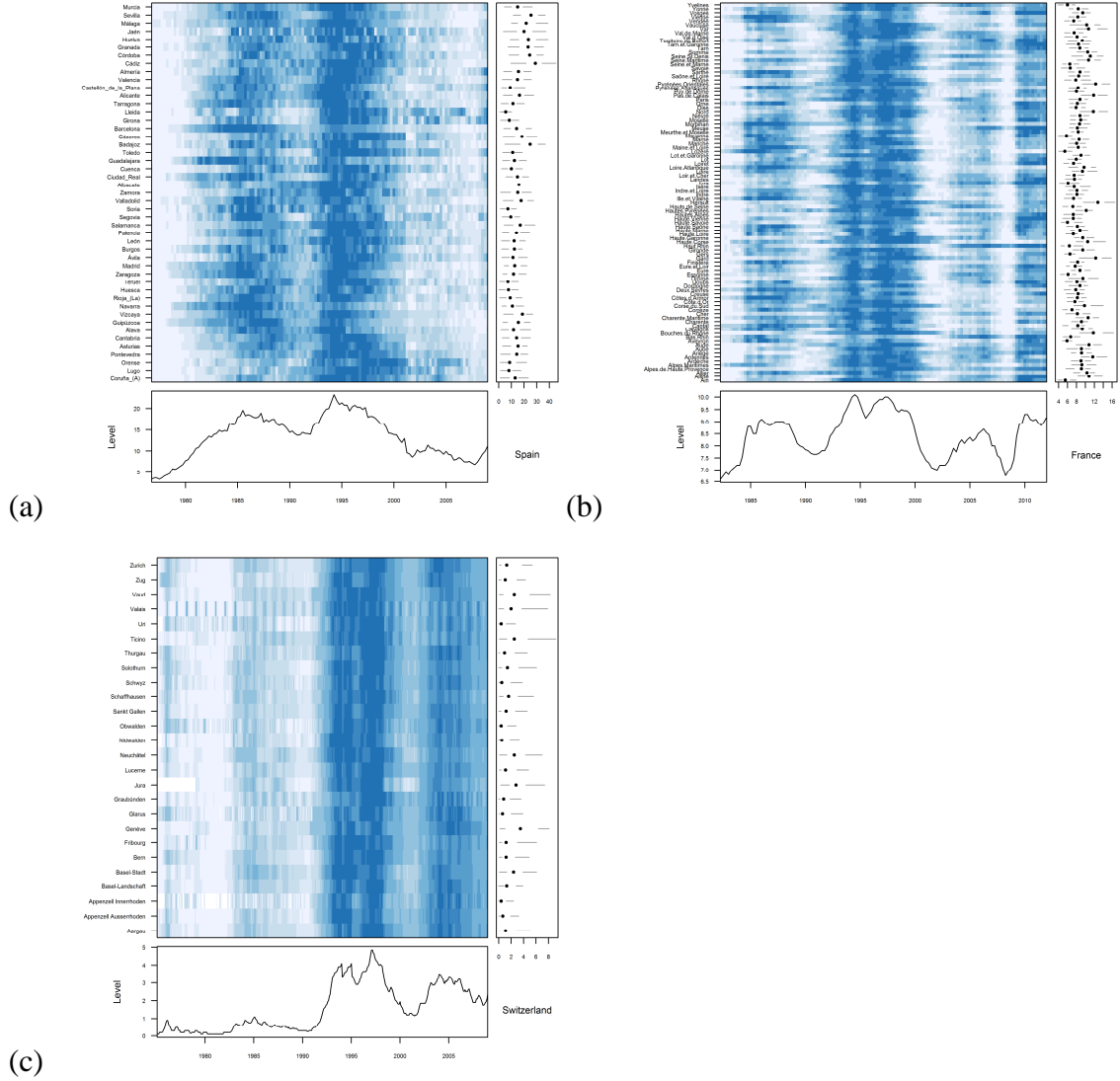inequalities have gradually built up over time, reaching a moderate level of spatial autocorrelation.

Figure 1. Panel plots of regional unemployment rates for Spain (a), France (b) and Switzerland (c)

3.2. Forecasting Strategy

Once the data peculiarities have been analysed, we now outline our forecasting strategy aiming to comparatively evaluate the desirability of the reviewed methods over different forecasting horizons using multi-step forecasts. The deterioration process of the forecasts is compared against differences in the spatial and temporal dimensions of the data sets and the geographical size of the spatial units.

We devise a forecasting strategy based on a rolling window and an extendable forecasting horizon. For each method and data set, estimates are obtained using a fixed-size window of observations (with the temporal index t being comprised between $1 + g$

9

and $T - h + g$), where $g = 0, ..., h - 1$, and h is the number of time periods covered by the one-year rolling window [i.e., $h = 4$ (quarters) for Spain and France, and $h = 12$ (months) for Switzerland]. *Ex post* forecasts of regional unemployment rates from one period (one-step-ahead) to a maximum of two years are then carried out. As a result, the number of forecasts obtained for each length of the forecasting horizon, and given cross-sectional dimensions, is $(4 * 47) = 188$ for Spain, $(4 * 96) = 384$ for France, and $(12 * 26) = 312$ for Switzerland.
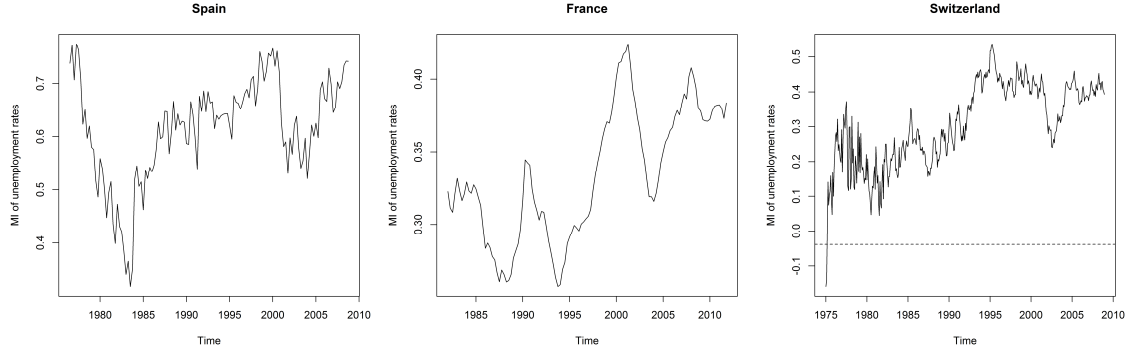

Figure 2. Dynamics of spatial autocorrelation, measured by MI, for Spain, France and Switzerland

3.3. Evaluation of Forecasts

The forecasting performance of our SVAR and SF models is compared by means of common statistical indicators: the mean squared error (MSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE). The MSE and MAE measure absolute deviations (i.e. regardless of direction) from the true values, and are computed as follows:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{U}_i - U_i\right)^2; \tag{5}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{U}_i - U_i\right|. \tag{6}$$

Because it is important to take into account scale heterogeneity in the levels of unemployment, we also consider the MAPE, which transforms the forecasting errors on a percentage scale, and is given by the following:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{\left|\hat{U}_i - U_i\right|}{U_i}\times 100. \tag{7}$$

Each model is tested on out-of-sample data for the years 2007–08 for Switzerland and Spain, and 2010–11 in the case of France. Within this framework, the discussion on the advantages and disadvantages of *ex ante* and *ex post* predictions seems unnecessary, since contemporaneous spatial lags are obtained in the first stage, as described in

Section 2. Angulo and Trívez (2010) avoid this debate as well when employing a dynamic panel data model in forecasting employment levels in Spanish provinces.

Starting from the MAPE (but the same could be done for MSE and MAE), we generate inferential results on the relative forecasting performance of our two competing models. We follow [anonymized] (2008) and [anonymized] (2012), and make use of the sign test (ST, Lehmann 1998). The ST is a nonparametric test for model comparison and aims to assess whether two models are equally accurate. It does not rely on the usual assumptions necessary for most alternative tests (such as the Diebold-Mariano test or the Wilcoxon test), since it does not require a normal distribution of the forecasting errors, or symmetry between the two vectors of errors compared. More simply, the ST compares the forecasting errors pairwise. If two models present a similar forecasting performance, the number of forecasts of Model 1 which show a greater error than the ones of Model 2 may be expected to be about 50 per cent of the total. Consequently, Model 1 is considered superior to Model 2 if Model 2 has higher forecasting errors in more than 50 per cent of the cases. In practical terms, a null hypothesis of equality in the median forecasting error is tested. The test statistic S is computed as:

$$S = \left(C - \frac{p}{2} \right) \bigg/ \frac{\sqrt{p}}{2},\tag{8}$$

where C is the number of times that Model 2 shows a higher error than Model 1's, and p is the total number of forecasts. The S statistic follows a normal distribution $N(0,1)$.

Alternatively, further parametric and nonparametric tests could be applied for model comparison. Among others, [anonymized] (2008) employ, in addition to the ST, a parametric test, that is, the Morgan-Granger-Newbold test, while [anonymized] (2007) use a further nonparametric test, namely the Friedman statistic.

4. Results

4.1. Spain

Table 2 reports summary statistics for the MSE, MAE and MAPE obtained for Spanish provinces over eight different forecasting horizons: from one quarter to eight quarters (two years). As mentioned in Section 3.2, the forecasting errors analysed for each forecasting horizon are the results of repeated estimations of the models using a one-year rolling window.

From the analysis of Table 2, the SVAR model appears to show better forecasting performance than the SF model, although the difference between the two models reduces slightly when MAPE is considered. Then the SVAR model shows a comparatively better performance when the focus is on error minimization on the scale of the dependent variable. This is explained by the higher level of variability of the SF model in comparison to the SVAR, as shown in Figure 3, where box plots provide a visualization of the error distribution. With regard to the main question of the paper, we observe – as expected – a deterioration of forecasting accuracy for longer forecast horizons, especially when $T = 5$ and more (i.e., beyond one year). The same conclusion is reached regardless of the error measure considered, although the SF model error appears to stabilize and actually decrease on average (and at the median) for longer

forecasts. Noteworthy is also the generalized increase in outlying forecasting errors, in particular when the forecasting horizon approaches two years and moves into the 2008 financial crisis, which had a strong impact on the Spanish labour market.

Table 2, Summary statistics of MSE, MAE and MAPE over expanding forecasting horizons (Spain)

Indicator	Model	1	2	3	4
MSE	SVAR	1.694	3.737	2.721	4.189
	SF	1.892	3.708	5.376	7.295
MAE	SVAR	0.999	1.404	1.275	1.567
	SF	1.081	1.539	1.907	2.248
MAPE	SVAR	0.122	0.163	0.154	0.181
	SF	0.136	0.204	0.263	0.323
ST (MAPE)	Winner	–	–	SVAR	SVAR
		5	6	7	8
MSE	SVAR	4.481	5.179	5.658	10.007
	SF	9.077	9.882	9.643	9.031
MAE	SVAR	1.649	1.760	1.718	2.308
	SF	2.257	2.690	2.624	2.479
MAPE	SVAR	0.203	0.292	0.342	0.384
	SF	0.373	0.389	0.363	0.316
ST (MAPE)	Winner	SVAR	SVAR	–	–

We compute the ST (on the basis of the MAPE indicator) to assess whether our two models can be considered as equally accurate for the Spanish data set. We calculate the test by pooling the forecasting errors by forecasting horizon. The test is not significant for the first two (shorter) horizons, then becomes significant, in favour of the SVAR, and finally, as observed in Figure 3, the SF model becomes competitive again at the two-year forecasting horizon.

If the existence of spatial patterns in the data and in the speed of response to shocks is well identified by both models, forecasting errors should not present spatial autocorrelation (unless, e.g., asymmetric shocks are faced). To test this hypothesis, MI is computed on the prediction errors of both methods for each iteration and forecasting horizon. It can be observed, in Figure 4, that the forecasting errors of the SF model show, in about 50 per cent of the cases, a significant and moderate spatial autocorrelation, while in most cases, SVAR errors are spatially uncorrelated.

4.2. France

In

Table 3 we report summary statistics for the eight forecasting horizons tested (i.e., as for Spain, up to two years) in the case of France. At first sight, the SF model appears to have gained in competitiveness from the different data structure (the unbalance between

n and *T* is now of a lesser extent). In particular, the SF model wins over the SVAR model for all forecasting horizons when MSE is considered (i.e., when the focus in on quadratic error). Additionally, the values for MSE are sensibly smaller than in the Spanish case (especially for the SF model), due to the limited variability of the French data (as suggested also for spatial autocorrelation in Figure 2).
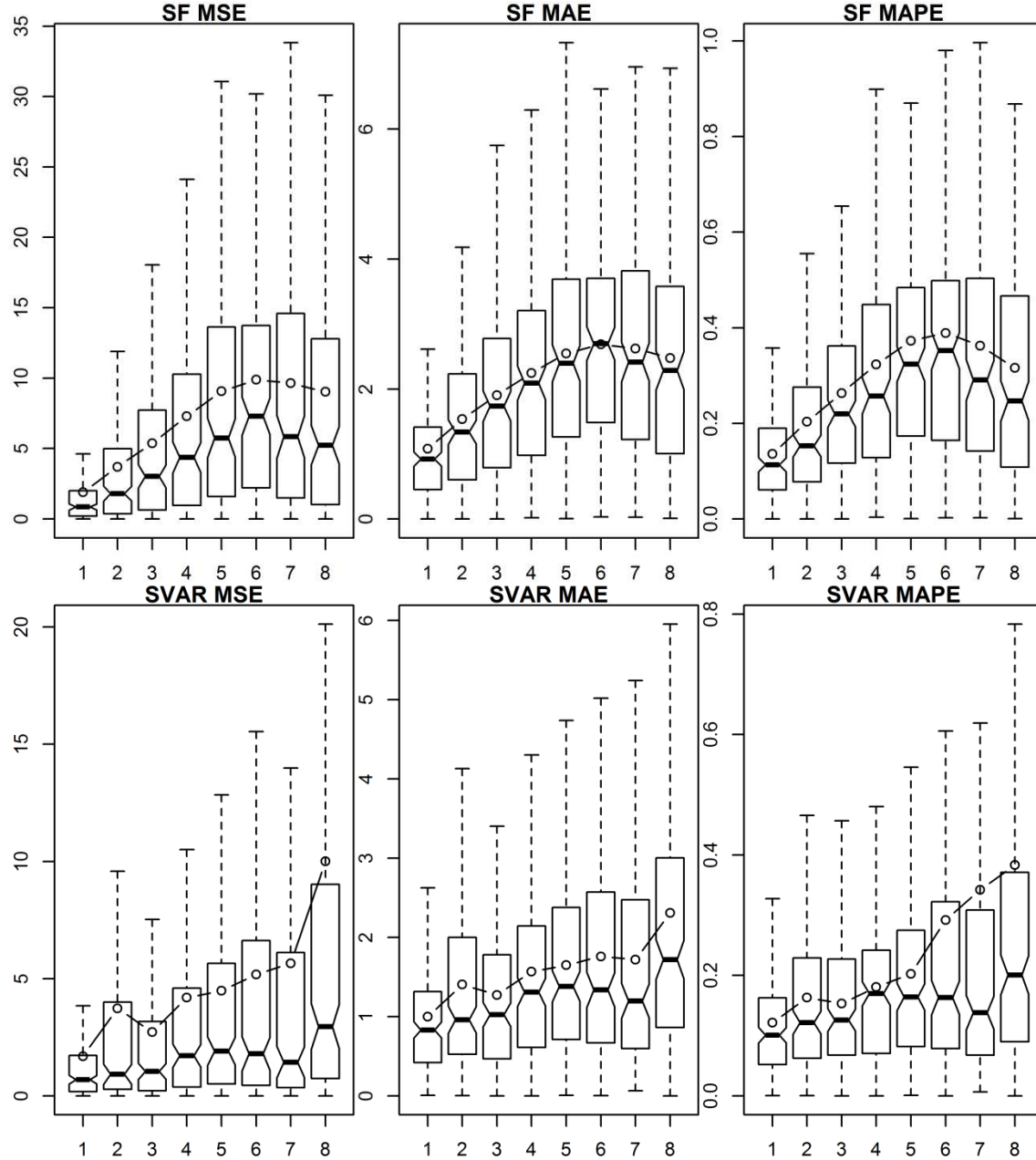

Figure 3. MSE, MAE and MAPE box plots over expanding forecasting horizons (Spain)

The SVAR model shows a better average performance only for relatively short forecasting horizons (*T* = 2–4), with regard to MAE and MAPE, pointing to a different forecast deterioration process of the models for forecasting horizons longer than one

year. The error indicators for the SF model do not increase over forecasting horizons (they actually decrease in some cases), while the SVAR model shows a monotonic forecast deterioration process. Additionally, Figure 5 shows that, while the SF model's outlying errors remain in the same range over more distant forecasts, the ones for the SVAR model increase greatly, leading also to a quick rising of the average error levels.
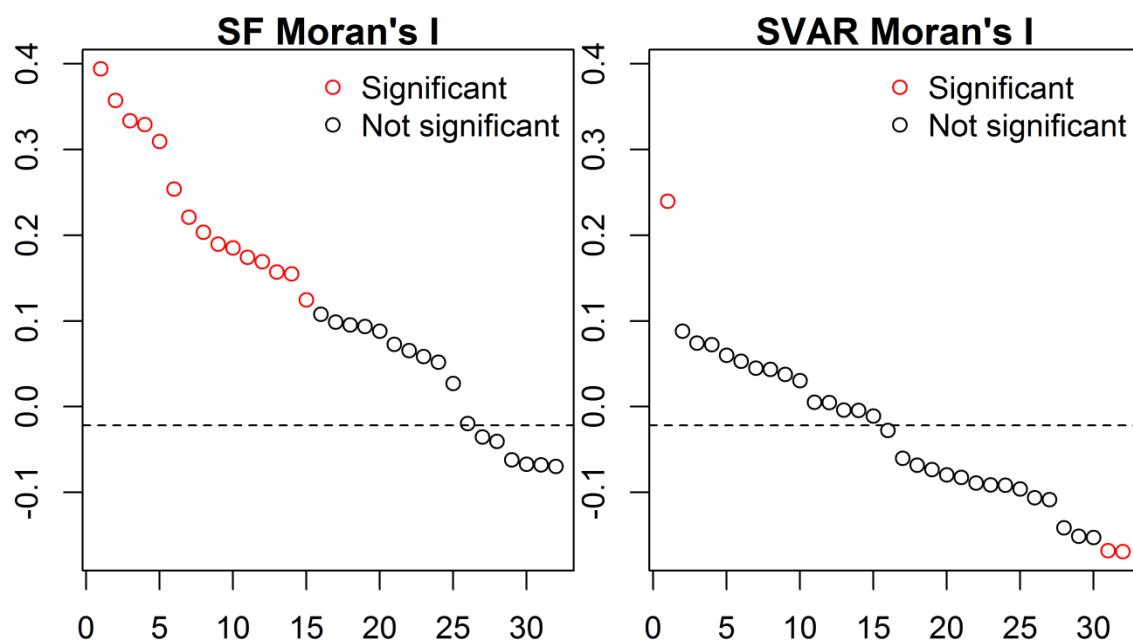

Figure 4. MI of SF and SVAR forecasting errors, sorted in decreasing order (Spain). Red dots identify significant values

Table 3. Summary statistics of MSE, MAE and MAPE over expanding forecasting horizons (France)

| Indicator	Model	1	2	3	4
MSE	SVAR	0.785	0.840	1.022	1.339
	SF	0.330	0.654	0.697	0.904
MAE	SVAR	0.502	0.511	0.564	0.641
	SF	0.471	0.625	0.656	0.772
MAPE	SVAR	0.054	0.056	0.062	0.070
	SF	0.053	0.068	0.072	0.076
ST (MAPE)	Winner	SVAR	SVAR	SVAR	SVAR
		5	6	7	8
MSE	SVAR	1.892	2.754	3.739	5.265
	SF	0.809	0.688	0.683	0.652
MAE	SVAR	0.775	0.947	1.098	1.299
	SF	0.670	0.643	0.649	0.635
MAPE	SVAR	0.086	0.109	0.135	0.188
	SF	0.073	0.071	0.072	0.070
ST (MAPE)	Winner	SVAR	–	SF	SF
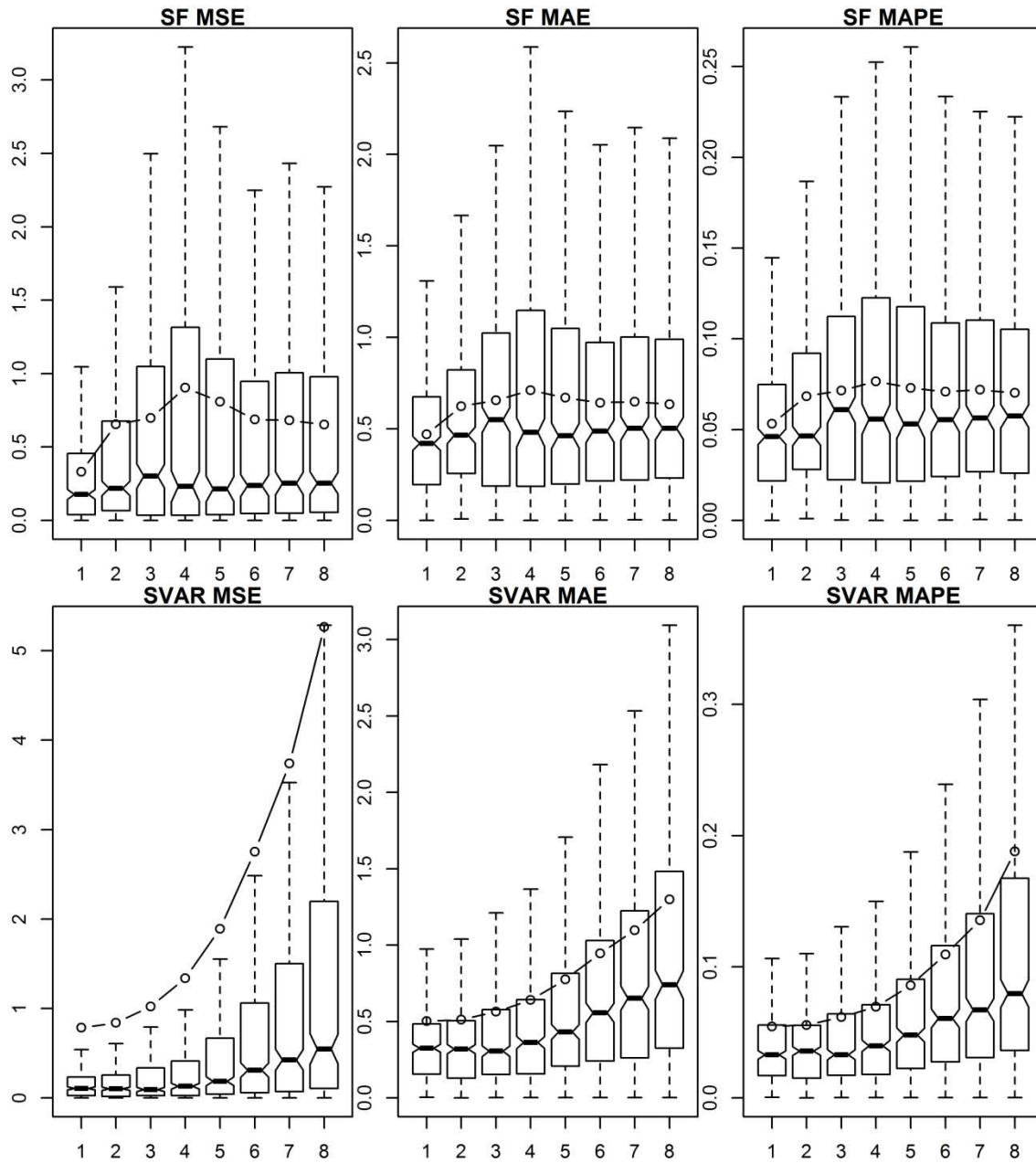
Figure 5. MSE, MAE and MAPE box plots over expanding forecasting horizons (*France*)

These findings, which are obtained for the average forecasting errors, are partly confirmed when the ST, which is carried out for the median error, is computed. The SVAR model appears to be significantly superior for short-term forecasts, while the SF model becomes competitive as the length of forecasting horizon increases, eventually emerging as significantly superior for forecasts seven or eight quarters ahead.

Finally, Figure 6 graphically summarizes our finding pertaining to the spatial autocorrelation of the SF and SVAR forecasting errors, leading to a conclusion similar to the one of the Spanish case. The two models appear to behave in opposite ways in

this regard, as the SF model once again produces forecasts with (moderately) spatially correlated errors, while the SVAR, though showing a smaller number of spatially autocorrelated errors, records a high number of cases in which forecasting errors are negatively and significantly autocorrelated. This issue, which could be related to the high error outliers discussed above or to core-periphery trends that the SVAR model cannot catch, deserves more specific attention in further work.
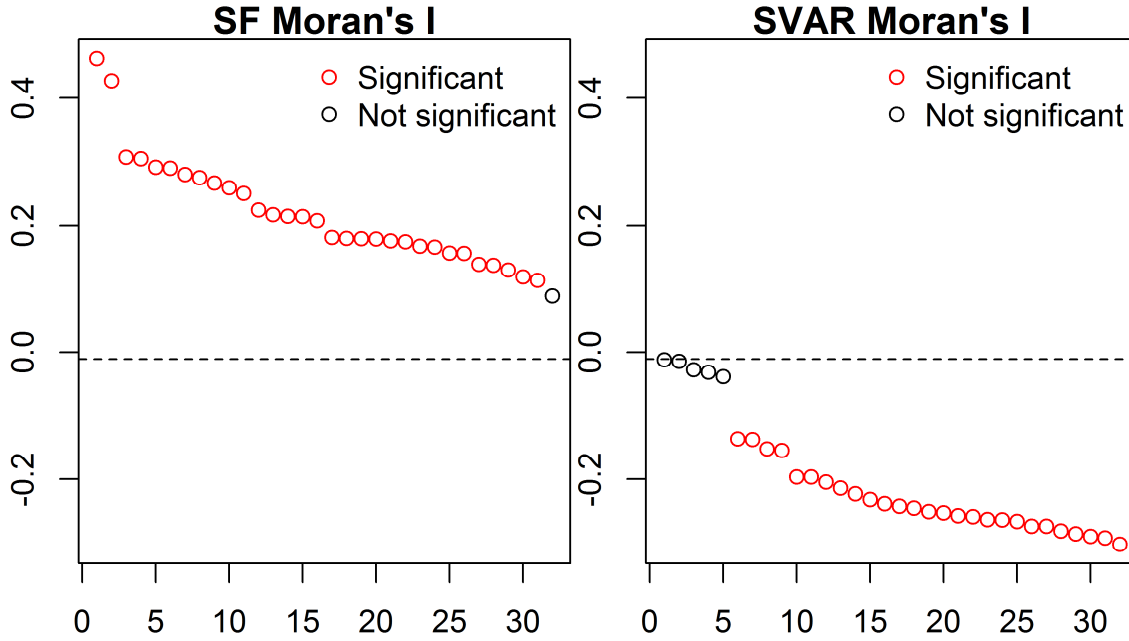

Figure 6. MI of SF and SVAR forecasting errors, sorted in decreasing order (France). Red dots identify significant values

4.3. Switzerland

Our final empirical exercise regards the 26 Swiss Cantons. The Swiss data set is, in theory, the one most in favour of the SVAR model, since the temporal dimension greatly exceeds the cross-sectional one ($T >> n$). The average forecasting errors reported in Table 4 confirm our expectation, as the SVAR model appears to produce smaller forecasting errors, on average, for all three indicators. The box plots appearing in Figure 7 visually confirm the average results, showing that for any forecasting horizon, median and average errors for the SVAR model, however computed, are smaller than the ones of the SF model.

The ST, computed on MAPE, comes as a further confirmation of what could be seen numerically and graphically, since the SVAR model emerges as producing significantly more accurate forecasts than the SF model. An exception to this finding are the two shortest forecasting horizons, where no significant distinction could be made between the two models.

Table 4. Summary statistics of MSE, MAE and MAPE over expanding forecasting horizons (Switzerland)

|		1	2	3	4	5	6	7	8	9	10	11	12
MSE	SVAR	0.016	0.034	0.044	0.053	0.060	0.063	0.065	0.074	0.081	0.082	0.082	0.086
	SF	0.020	0.054	0.090	0.124	0.151	0.178	0.204	0.227	0.250	0.271	0.291	0.319
MAE	SVAR	0.092	0.138	0.162	0.179	0.192	0.202	0.208	0.223	0.232	0.232	0.231	0.234
	SF	0.102	0.170	0.221	0.267	0.301	0.340	0.380	0.417	0.449	0.474	0.498	0.525
MAPE	SVAR	0.040	0.060	0.070	0.078	0.085	0.090	0.093	0.100	0.106	0.107	0.108	0.110
	SF	0.043	0.071	0.092	0.112	0.127	0.145	0.167	0.185	0.205	0.220	0.234	0.247
ST (MAPE)	Winner	–	–	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR
		13	14	15	16	17	18	19	20	21	22	23	24
MSE	SVAR	0.097	0.112	0.114	0.119	0.123	0.112	0.109	0.111	0.110	0.110	0.111	0.118
	SF	0.375	0.444	0.511	0.568	0.612	0.648	0.672	0.692	0.706	0.705	0.696	0.689
MAE	SVAR	0.247	0.262	0.268	0.271	0.273	0.263	0.258	0.262	0.261	0.259	0.257	0.265
	SF	0.566	0.613	0.653	0.687	0.712	0.732	0.748	0.761	0.768	0.767	0.761	0.753
MAPE	SVAR	0.116	0.124	0.128	0.132	0.133	0.129	0.126	0.127	0.126	0.126	0.125	0.128
	SF	0.269	0.294	0.318	0.338	0.353	0.365	0.376	0.387	0.392	0.392	0.389	0.384
ST (MAPE)	Winner	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR	SVAR
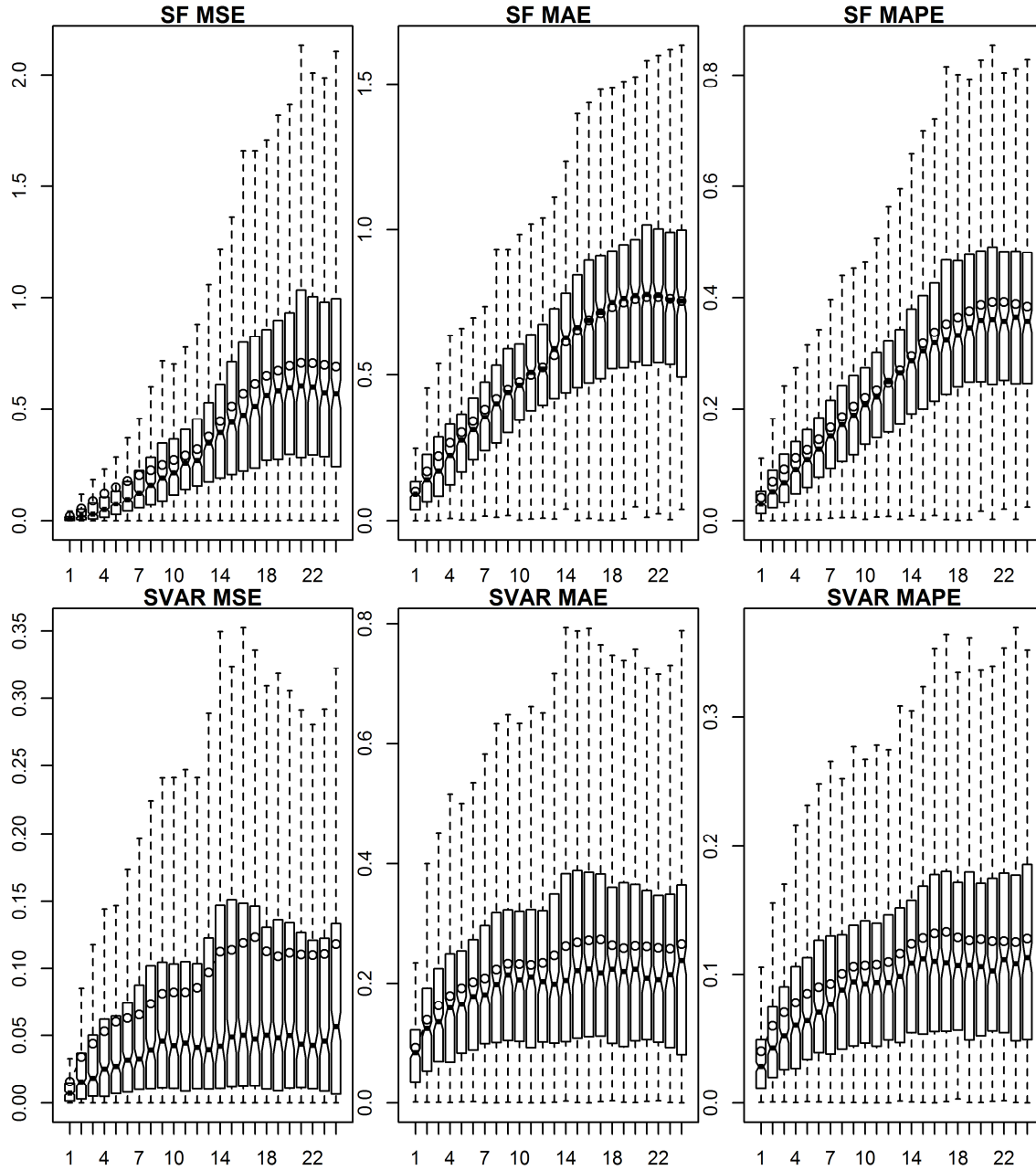
Figure 7. MSE, MAE and MAPE box plots over expanding forecasting horizons (Switzerland)

Finally, Figure 8 reports, in descending order, the values of MI for the level of spatial autocorrelation of forecasting errors. As seen above for Spain and France, the SF method implies some positive (moderate) spatial autocorrelation which is instead more limited (i.e., there are less statistically significant cases) for the SVAR. This recurring finding may be due to the imposition of a time-invariant spatial structure (the spatial filter) at the level of the intercept (i.e., a similar result could be expected if using conventional individual fixed effects), which may partially constrain regional trends within a defined spatial pattern. This issue calls for further evaluation, for example for the case of asymmetric shocks that may affect only a certain group of regions or specific sectors in which only some areas are specialized.
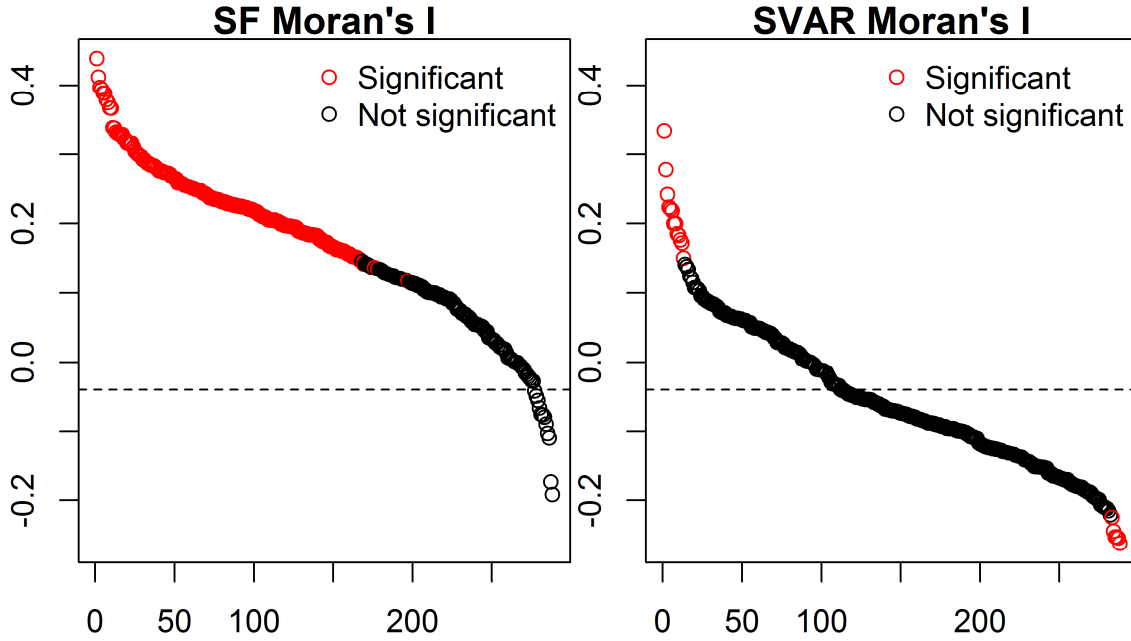
Figure 8. MI of SF and SVAR forecasting errors, sorted in decreasing order (Switzerland). Red dots identify significant values

5. Conclusions

In this paper we evaluated two competing spatial (dynamic) methods for panel data forecasting: a spatial vector-autoregressive model (SVAR) and a dynamic panel data model employing spatial filtering (SF). For this purpose, we selected different real-world data sets (for Spain, France and Switzerland), all dealing with regional unemployment rates and characterized by different cross-sectional and temporal dimensions. Our main objective was to analyse the process of deterioration of forecasting precision as the forecasting horizon increases. In this view, we carried out a sensitivity analysis testing how different cross-sectional and temporal dimensions (i.e., the number of spatial units and of observations available per unit, respectively) influence the models' relative forecasting performance.

After empirical evaluation, the aforementioned differences in data structure indeed appear to be a discriminating factor in terms of forecasting accuracy. The SVAR model seems to be preferred to the SF model – in terms of both average and median error – when the temporal dimension is much greater than the spatial dimension and the spatial units have smaller size and a greater degree of variability (i.e., the Swiss data). Once the balance between n and T becomes fairer (partly in the data for Spain and more prominently in those for France, where T is only 1.25 times n), our results tend to become less uniform. In this regard, the length of the forecasting horizon becomes a diriment factor: while the SVAR model is generally outperforming the SF model for short-run forecasts, the SF model gains competitiveness – eventually emerging as the preferable model – for more distant forecasts.

Finally, we investigated whether the forecasting errors of the SVAR and SF models exhibited any spatial pattern, testing for spatial autocorrelation. The SVAR model

showed a smaller number of cases of spatially autocorrelated errors than the SF model, for both the Swiss and the Spanish data sets. Less straightforward results were found for the French data set, for which both positive and negative spatial autocorrelation was found.

Ultimately, our findings suggest that both methods (SVAR and SF) may deserve their own niche in regional forecasting. On the one hand, time-series-based spatial methods such as SVAR models emerge, as expected, if some depth in the temporal dimension is present, so to exploit the region-specific information on adjustment speed. On the other hand, spatial methods focusing more on the relevance of the cross-sectional dimension, like SF-augmented dynamic panel data models, gain in competitiveness in empirical cases when there is a fair balance between the cross-sectional and temporal dimensions.

While our results are not qualitatively surprising, they call for more questions to be answered. In particular with regard to our findings, one may wonder to what extent are the cross-sectional and temporal dimensions actually influencing our results, and to what extent are instead the geographical characteristics of the regions (e.g. their area) or macro-attributes (institutional differences at the national level influencing labour mobility) that drive our indicators. It may be worthwhile, to investigate this issue, to numerically evaluate, within a regression framework, how per-region forecasting error depends on all of the above factors.

Additionally, it would be desirable to expand our forecasting exercise to a greater number of data sets, comprising the cases of small-n small-T (e.g., Italian data would provide such a setting, though at the NUTS-2 level of geographical aggregation) and big-n small-T (as mentioned in Section 3, German data would fit this format), though the latter would imply imposing restrictions in order to estimate the SVAR model. The addition of further estimation approaches would fittingly complement our analysis.

A further issue of potential interest is the one pertaining to the level of spatial autocorrelation of forecasting errors. Our findings suggest that the SVAR model does a better job in this regard, although we did not delve into a detailed analysis of how such spatial autocorrelation distributes over different forecasting horizons. If the SF model, as it appears, improves its comparative performance for longer horizons, the spatial autocorrelation of the related forecasting errors might well follow the same pattern. We leave this and the above questions to further research.

References

Angulo AM, Trívez FJ (2010) The Impact of Spatial Elements on the Forecasting of Spanish Labour Series. Journal of Geographical Systems 12 (2):155-74

Bai J (2013) Fixed-Effects Dynamic Panel Models, a Factor Analytical Method. Econometrica 81 (1):285-314

Baltagi BH, Bresson G, Pirotte A (2012) Forecasting with Spatial Panel Data. Computational Statistics and Data Analysis 56 (11):3381-97

Baltagi BH, Li D (2004) Prediction in Panel Data Model with Spatial Correlation. In: Anselin L, Florax RJGM, Rey SJ (eds) Advances in Spatial Econometrics: Methodology, Tools and Application. Springer-Verlag, Berlin Heidelberg, pp. 283-95

Baltagi BH, Li D (2006) Prediction in the Panel Data Model with Spatial Correlation: the Case of Liquor. Spatial Economic Analysis 1 (2):175-85

Barca F, McCann P, Rodríguez-Pose A (2012) The Case for Regional Development Intervention: Place-Based versus Place-Neutral Approaches. Journal of Regional Science 52 (1):134-52

Beenstock M, Felsenstein D (2007) Spatial Vector Autoregressions. Spatial Economic Analysis 2 (2):167-96

Blanchard O, Jimeno JF (1995) Structural Unemployment: Spain versus Portugal. The American Economic Review 85 (2):212-8

Blanchard OJ, Katz LF (1992) Regional Evolutions. Brookings Papers on Economic Activity 1:1-75

Blundell R, Bond S (1998) Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. Journal of Econometrics 87 (1):115-43

Bun MJG, Carree MA (2005) Bias-Corrected Estimation in Dynamic Panel Data Models. Journal of Business and Economic Statistics 23 (2):200-10

Canova F, Ciccarelli M (2013) Panel Vector Autoregressive Models: A Survey. Working Paper Series 1507, European Central Bank, Frankfurt

Di Giacinto V (2003) Differential Regional Effects of Monetary Policy: A Geographical SVAR Approach. International Regional Science Review 26 (3):313-41

Fingleton B (2009) Prediction Using Panel Data Regression with Spatial Random Effects. International Regional Science Review 32 (2):195-220

Fotheringham AS, Brunsdon C, Charlton M (2002) Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. John Wiley & Sons Ltd, Chicester

Giacomini R, Granger CWJ (2004) Aggregation of Space-Time Processes. Journal of Econometrics 118 (1–2):7-26

Griffith DA (2000) A Linear Regression Solution to the Spatial Autocorrelation Problem. Journal of Geographical Systems 2 (2):141-56

Griffith DA (2003) Spatial Autocorrelation and Spatial Filtering: Gaining Understanding through Theory and Scientific Visualization. Springer-Verlag, Berlin Heidelberg New York

Griffith DA (2008) Spatial-Filtering-Based Contributions to a Critique of Geographically Weighted Regression (GWR). Environment and Planning A 40 (11):2751-69

Hernández-Murillo R, Owyang MT (2006) The Information Content of Regional Employment Data for Forecasting Aggregate Conditions. Economics Letters 90 (3):335-9

Isiklar G, Lahiri K (2007) How Far Ahead Can We Forecast? Evidence from Cross-Country Surveys. International Journal of Forecasting 23 (2):167-87

Jimeno JF, Bentolila S (1998) Regional Unemployment Persistence (Spain, 1976-1994). Labour Economics 5 (1):25-51

Kholodilin KA, Mense A (2012) Forecasting the Prices and Rents for Flats in Large German Cities. 1207, DIW, Berlin

Kholodilin KA, Siliverstovs B, Kooths S (2008) A Dynamic Panel Data Approach to the Forecasting of the GDP of German Länder. Spatial Economic Analysis 3 (2):195-207

Lehmann EL (1998) Nonparametrics: Statistical Methods Based on Ranks. rev. edn. Prentice Hall, Upper Saddle River

Longhi S, Nijkamp P (2007) Forecasting Regional Market Developments under Spatial Autocorrelation. International Regional Science Review 30 (2):100-51

Moran P (1948) The Interpretation of Statistical Maps. Journal of the Royal Statistical Society B 10:243-51

Ohtsuka Y, Kakamu K (2013) Space-Time Model versus VAR Model: Forecasting Electricity Demand in Japan. Journal of Forecasting 32 (1):75-85

Pan Z, LeSage JP (1995) Using Spatial Contiguity as Prior Information in Vector Autoregressive Models. Economics Letters 47 (2):137-42

Partridge MD, Rickman DS (1995) Differences in state unemployment rates: the role of labor and product market structural shifts. Southern Economic Journal 62 (1):89-106

Patacchini E, Zenou Y (2007) Spatial Dependence in Local Unemployment Rates. Journal of Economic Geography 7 (2):169-91

Pesaran MH, Schuermann T, Weiner SM (2004) Modeling Regional Interdependencies Using a Global Error-Correcting Macroeconometric Model. Journal of Business and Economic Statistics 22 (2):129-62

Schanne N, Wapler R, Weyh A (2010) Regional Unemployment Forecasts with Spatial Interdependencies. International Journal of Forecasting 26 (4):908-26

Sims CA (1980) Macroeconomics and Reality. Econometrica 48 (1):1-48