Modeling Spatiotemporal Polarization of Resident Foreign Population

Giampiero Marra, David L. Miller, and Luca Zanin

May 10, 2010

1 Introduction

During the last decades, many scholars have shown an increasing interest in finding an explanation to the phenomenon of international migrations, as well as in monitoring the process of integration of foreigners in host countries. Several theories have been proposed, from both a macro and micro point of view, on the basis of different lines of thought. We mention, for example, the 'neoclassical economic theory' (e.g., Borjas 1989; Massey et al. 1993), the recent assumptions of the 'new economics of migration' (e.g., Stark and Bloom 1985; Massey et al. 1993), the aspects of 'dual labor market' (e.g., Piore 1979; Massey et al. 1993) and the 'world systems theory' (e.g., Wallerstein 1974; Portes and Walton 1981; Morawska 1990; Hooghe et al. 2008). In supporting international movements are also rather important 'migrant networks' whose role is to connect migrants and non-migrants in origin and destination areas through ties of friendship, shared community origin, etc. (e.g., Massey et al. 1993). A large number of empirical cases have also been analyzed. A non-exhaustive list of examples includes the studies that have investigated the impact of foreigners on the labor market of the host country (e.g., Borjas 1994, 2003, 2005; Card 2005; Fullin and Reyneri 2010), the aspect of underemployment (Slack and Jensen 2007), the possible nexus between immigration and nation's poverty rate (Raphael and Smolensky 2009), the transmission of identities from the first to second generation immigrants (Casey and Dustmann 2010), the difference in terms of education, earnings and employment between first and second generation immigrants (Algan et al. 2010), the problems of integration of different cultures and languages (e.g., Lazer 1999; Contucci and Ghirlanda 2007), and the tendency of immigrants for living in ethnic "enclaves" (Edin et al. 2003). Despite the increasing literature on international migration, to the best of our knowledge, no studies have attempted to model the spatiotemporal polarization of the incidence of resident foreigners in the total population of a country, and this is the aim of this article. The incidence can be calculated as follows

$$\% \ incidence \ of \ resident \ foreigners_j = \frac{number \ of \ resident \ foreigners_j}{total \ resident \ population_j} \times 100 \qquad (1)$$

where j represents a specific area such municipality, province, or region.

In the special issue of *The Economic Journal* dedicated to 'the integration of immigrants and its consequences', Manning (2010) has pointed out as many European countries are experiencing a significant increase of immigrants in the population. As an example, according to the official statistics of the Italian National Statistical Office (ISTAT), in the period $2003 - 2008$, the incidence of resident foreigners in the

total population has increased from 3.4% to 6.5%. The speed of change in the indicator is remarkable if we recall that, on the basis of the population census, in the period $1991 - 2001$ the incidence has increased from 0.6% to 2.3%. The progressive growth of migration flows from abroad has been contributing to important changes from the socio-demographic and economic point of view; for instance, during the last years, the official statistics have highlighted an increasing frequency in the intermarriage between Italians and foreigners, as well as in the number of entrepreneurs with foreign citizenship. For the next years, the scenario up to 2050 produced by ISTAT suggests a progressive and substantial increase of resident foreigners (about 17% of the total population by 2050 as shown in Figure 2), a tendency that is common to many European countries (**REF**). Therefore, it is of paramount importance to formulate adequate social policies which can support, for instance, the integration process of immigrants in the host country. In this respect, it is crucial to promote the recognition and respect for different cultural identities with the aim of ensuring social cohesion between foreign and native people, and avoiding social unrest (e.g., Cheong et al. 2007). Immigrants' rights and integration policies are typically governed by specific laws which regulate pubblic health and education services (e.g., Zincone 2006), but also professional training and the right to housing, to name a few. Several policies dedicated to foreigners also focus on offering guidance services and counseling on legal issues and administrative procedures. In this context, policy-makers can be interested in knowing whether there are areas of the country in which such policies should be strengthened as a result of a higher presence of foreigners.
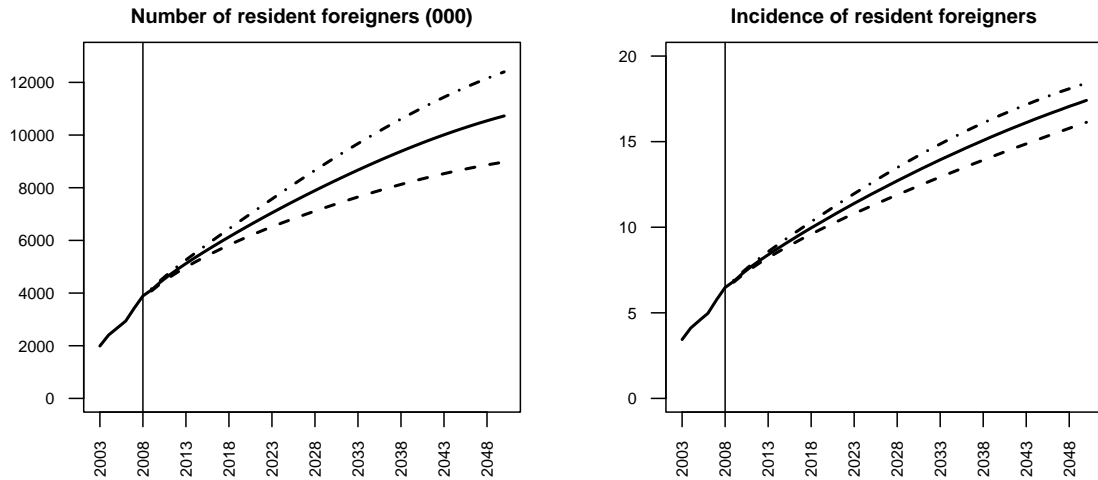

Figure 1: The graphs show the trend in the presence of foreigners resident in Italy during $2003 - 2008$, and in forecast for the period $2009 - 2050$, based on the official statistics and elaborations of ISTAT. The information is reported in terms of number (divided by 1000) and incidence of resident foreigners. The average, low and high scenarios are by the solid and dotted lines, respectively.

Measure (1) represents a simple well known demographic indicator for comparing different zones of a country in terms of number of foreigners every 100 resident inhabitants (e.g., Bauer and Zimmermann 1999; De Bartolo 2007; Lowell 2007). The indicator returns, in this sense, an information on the type of spatial distribution of foreigners (balanced/homogeneous or unbalanced/inhomogeneous), hence identifying areas of the country with a greater or lesser presence of resident foreigners in the total population. If such distribution is unbalanced in space and time, what are the factors that contribute to attract a larger number of foreigners in an area as compared to another? Inhomogeneous spatial incidence may be explained by several factors

related to the characteristics of a specific area such as the presence of foreigners with the same origin (recall, for example, the network theory mentioned earlier on), and the presence of dwellings available at more favorable prices than those in other areas. Here, we focus on factors of economic nature. Specifically, we construct a composite indicator that we have named 'Index of Spatial Economic Attractiveness'. It is made up of three relevant indicators, namely (i) added value per capita, (ii) unemployment rate and (iii) employment rate, and can be easily constructed for many countries given the use of standard and well known economic variables.

The empirical analysis proposed in this article addresses the issues mentioned in the previous paragraphs. Our study is carried out for Italy by using data from ISTAT which, since 2003, provides a public database with annual frequency on the number of resident foreigners at municipal level. This information is new in the panorama of Italian statistics, and currently represents the maximum spatial and temporal detail available. Spatiotemporal modelling is achieved via the use of splines which are a popular way of performing smoothing over some bounded domain $\Omega \subseteq \mathbb{R}^2$, over time (e.g., Hastie and Tibshirani 1990; Ruppert, Wand, and Carroll 2003; Wood 2006). In our context, the response (incidence of resident foreigners) is modelled as a function of its spatial coordinates and time trend. The estimated function can then be used to create smoothed maps of the geographical area of interest over time. When a geographical region has complex boundaries, as in many applied contexts, features from one part of the domain can unduly influence other parts giving rise to a phenomenon known as *leakage* (e.g., Ramsay 2002). It typically occurs when a smoother inappropriately links two pats of a domain. This is problematic since it causes the fitted surface to be mis-estimated, hence leading to incorrect inference (e.g., biased incidence estimates), which is clearly not desirable. Such an issue can be overcome if the proposed spatiotemporal model can account for the structure of the domain that is under investigation. In practice, this can be achieved by using several approaches (Ramsay 2002, and references therein). Here, we construct our model by using a three-dimensional tensor product smoother combining a cubic regression spline basis function for time trend (e.g., Wood 2006) and a soap film spline basis for the spatial dimensions (Wood et al. 2008).

The remainder of the article is organized as follows.

2 Data

2.1 Sources, definitions and critical issues in the statistics of resident foreigners

LUCA describes datasource etc.

The incidence of resident foreigners exhibits a positively skewed distribution. For this reason, a gamma distribution provides a realistic description of the response variable. Notice that, during the model fitting process, all gamma distributions for all years will be assumed to have a common shape parameter since preliminary descriptive analysis suggested that this is the case.

2.2 Resident foreigners: the main features

LUCA explains the problem in Italy and son on.

3 Methodology

3.1 Model specification

The model we employ belongs to the class of generalized additive models (GAMs, Hastie and Tibshirani 1990). They allow for complex relationships between covariates and response variable, which are crucial to uncover interesting features in the data. The proposed model is as follows

$$\log\{\mathbb{E}(\texttt{inc.rf}_{it})\} = f(\texttt{year}_t, \texttt{n}_i, \texttt{e}_i), \quad \texttt{inc.rf}_{it} \sim \text{Gamma}, \tag{2}$$

for $i = 1, \ldots, 8094$ and $t = 1, \ldots, 6$. The log link function ensures positive fitted values. $\texttt{inc.rf}_{it}$, $\texttt{n}_i$, and $\texttt{e}_i$ represent the variables percentage incidence of resident foreigners, northing, and easting, respectively. The function f is a multidimensional smooth of $\texttt{year}$, $\texttt{n}$, and $\texttt{e}$ which models the joint effect of these variables on $\texttt{inc.rf}$. Notice that we want both the space and time dimension to have an optimal degree of smoothness in terms of the bias variance trade-off. This means that the chosen smoother has to be invariant to the relative scaling of space (km) and time (years). This can be achieved by using a multidimensional tensor product smooth combining a cubic regression spline basis for time trend and an *isotropic* soap film spline basis for the two spatial dimensions $\texttt{n}$ and $\texttt{e}$, since the smoother should not depend on the coordinate system used and account for the structure of the geographical domain under investigation (details are given in the next two sections). The smooth component in (2) is subject to identifiability constraints; see Wood (2006) for more details.

3.2 A three-dimensional tensor product smoother for time and space

The construction of a three-dimensional scale invariant tensor product smoother of time and space is based on a marginal one-dimensional spline basis for time and a two-dimensional marginal smooth for space, with associated quadratic penalties measuring their roughness. We omit the subscripts i and t for simplicity. Let us assume that we have two low-rank regression spline bases of any type to represent the smooth functions f_{year} and f_{space}, we can write (e.g., Ruppert, Wand, and Carroll 2003)

$$f_{year}(\texttt{year}) = \sum_{l=1}^{L} \alpha_l a_l(\texttt{year}) = \mathbf{X}_{year}\boldsymbol{\alpha} \quad \text{and} \quad f_{space}(\texttt{n}, \texttt{e}) = \sum_{r=1}^{R} \gamma_r d_r(\texttt{n}, \texttt{e}) = \mathbf{X}_{space}\boldsymbol{\gamma},$$

where the $a_l(\texttt{year})$ and $d_r(\texttt{n}, \texttt{e})$ are known cubic regression spline and soap film basis functions, with corresponding parameters α_l and γ_r, L and R are the spline dimensions of the two smooth components, and $\mathbf{X}_{year}$ and $\mathbf{X}_{space}$ are marginal model matrices evaluating the basis functions with parameter vectors $\boldsymbol{\alpha}$ and $\boldsymbol{\gamma}$ (Wood 2006, 2008). In order to set up a three-dimensional tensor product smoother for time and space we need $f_{year}(\texttt{year})$ to vary smoothly within the space dimensions. This can be achieved by allowing the parameters α_l to vary smoothly with $\texttt{n}$ and $\texttt{e}$. Using the spline set-up for $f_{space}(\texttt{n}, \texttt{e})$ we can write (e.g., Wood 2006)

$$\alpha_l(\texttt{n}, \texttt{e}) = \sum_{r=1}^{R} \gamma_{lr} d_r(\texttt{n}, \texttt{e})$$

which results in

$$f(\texttt{year}, \texttt{n}, \texttt{e}) = \sum_{l=1}^{L} \sum_{r=1}^{R} \gamma_{lr} d_r(\texttt{n}, \texttt{e}) a_l(\texttt{year}).$$

For any particular set of observations of $\texttt{year}$, $\texttt{n}$, and $\texttt{e}$, there exists a simple relationship between the matrix $\mathbf{X}$ evaluating the tensor product smooth at these observations, and the model matrices $\mathbf{X}_{year}$ and $\mathbf{X}_{space}$ evaluating the marginal smooths at the same observations. Ordering appropriately the parameters γ_{lr} into a vector $\boldsymbol{\theta}$, the ith row of $\mathbf{X}$ is given by $\mathbf{X}_i = \mathbf{X}_{year,i} \otimes \mathbf{X}_{space,i}$, where $\otimes$ is the Kronecker product.

Within the GAM context, it is necessary to quantify the roughness of the smooth functions in the model so that over-fitting can be accounted for and hence avoided during the parameter estimation process (e.g., Marra and Radice 2010). As for the penalty associated with this tensor product basis, it is possible to start from roughness measures associated with the marginal smooths $f_{year}(\texttt{year})$ and $f_{space}(\texttt{n}, \texttt{e})$. Suppose that functionals Js measuring the roughness of the smooth terms are available, and that these can be written as quadratic forms in the marginal parameters, we have that

$$J_{year}(f_{year}) = \boldsymbol{\alpha}^{\mathsf{T}} \mathbf{S}_{year} \boldsymbol{\alpha} \quad \text{and} \quad J_{space}(f_{space}) = \boldsymbol{\gamma}^{\mathsf{T}} \mathbf{S}_{space} \boldsymbol{\gamma},$$

where the $\mathbf{S}$ matrices contain known coefficients whose values depend on the chosen bases for time and space. For instance, the second-order cubic spline penalty for $f_{year}(\texttt{year})$ evaluates $J_{year}(f_{year}) = \int \left(\partial^2 f_{year} / \partial year^2 \right)^2 dyear$, but it may be more complex (e.g., Wood 2006, 2008). Following, for example, Zanin and Marra (2010), an overall penalty for the tensor product smooth can be obtained by applying the penalties of $f_{space}(\texttt{n}, \texttt{e})$ to the varying coefficients of the marginal smooth $f_{year}(\texttt{year})$, $\alpha_l(\texttt{n}, \texttt{e})$,

$$\sum_{l=1}^{L} J_{space} \left\{ \alpha_l(\texttt{n}, \texttt{e}) \right\},$$

and the penalties of $f_{year}(\texttt{year})$ to the varying coefficients of the marginal smooth $f_{space}(\texttt{n}, \texttt{e})$, $\gamma_r(\texttt{year})$,

$$\sum_{r=1}^{R} J_{year} \left\{ \gamma_r(\texttt{year}) \right\}.$$

It follows that the roughness penalty of $f(\texttt{year}, \texttt{n}, \texttt{e})$ can be measured as

$$J(f) = \lambda_{space} \sum_{l=1}^{L} J_{space} \left\{ \alpha_l(\texttt{n}, \texttt{e}) \right\} + \lambda_{year} \sum_{r=1}^{R} J_{year} \left\{ \gamma_r(\texttt{year}) \right\},$$

which can also be written as

$$\lambda_{space} \boldsymbol{\theta}^{\mathsf{T}} \mathbf{I}_L \otimes \mathbf{S}_{space} \boldsymbol{\theta} + \lambda_{year} \boldsymbol{\theta}^{\mathsf{T}} \mathbf{S}_{year} \otimes \mathbf{I}_R \boldsymbol{\theta},$$

where, once again, the vector $\boldsymbol{\theta}$ contains the tensor product smooth parameters. The λ are smoothing parameters controlling the trade-off between model fit and model smoothness. The next section shows how f_{space} and $\mathbf{S}_{space}$ can be constructed.

3.2.1 Soap film

DAVE will start writing this part, which will be then checked with GIAMPIERO.

3.3 Parameter estimation

In model (2), replacing f with its tensor product expression yields a generalized linear model (McCullagh and Nelder, 1989) whose design matrix contains the spline bases representing the smooth component in the model. This means that in principle such a model can simply be estimated by maximum likelihood (ML). However, in a smoothing spline context, unpenalized parameter estimation is likely to result in smooth component estimates that are too 'wiggly', hence undermining the utility of such a model. This can be overcome by penalized ML, where the use of penalties allows for the suppression of that part of smooth term complexity which has no support from the data (e.g., Marra and Radice 2010). Specifically, the model can be fitted by minimization of

$$D(\boldsymbol{\theta}) + \boldsymbol{\theta}^{\mathsf{T}} \mathbf{S} \boldsymbol{\theta} \quad \text{w.r.t.} \quad \boldsymbol{\theta}. \tag{3}$$

The model deviance, D, is defined as $2\phi(l_{\text{sat}} - l)$, where ϕ is a dispersion parameter, l is the log-likelihood of the model and l_{sat} the maximum value for the log-likelihood of the model with one parameter per datum. $\mathbf{S} = \sum_i \lambda_i \mathbf{S}_i$, where the $\mathbf{S}_i$ are matrices of known coefficients properly defined according to the results of the previous sections, and the λ_i are smoothing parameters associated to them.

Given values for the λ_i, minimization of (3) is straightforward. However, smoothing parameter estimation has to be addressed. This can be achieved by minimization of a prediction error estimate, such as the generalized cross-validation (GCV) (Craven and Wahba 1979), or by approximate restricted ML (REML) estimation (Wood 2010). Smoothing parameter selection via the GCV consists of minimizing

$$V(\boldsymbol{\lambda}) = \frac{nD(\hat{\boldsymbol{\theta}})}{\{n - \text{tr}(\mathbf{A})\}^2}, \tag{4}$$

where $\mathbf{A}$ is the usual hat matrix, and $\text{tr}(\mathbf{A})$ represents the effective degrees of freedom (edf) or number of estimated parameters of the penalized model. Wood (2006) described a computational procedure to estimate smoothing parameters on the basis of criterion (4). As an alternative, approximate REML can be employed. Within this framework, the penalized likelihood estimates, $\hat{\boldsymbol{\theta}}$, can be seen as the posterior modes of the distribution of $\boldsymbol{\theta}|\mathbf{y}$ if $\boldsymbol{\theta} \sim N(\mathbf{0}, \mathbf{S}^{-1})$, where $\mathbf{S}^{-1}$ is an appropriate generalized inverse. Viewing the spline parameters as random effects allows for the possibility to estimate the λ_i via REML (Wahba 1985). The recent work by Reiss and Ogden (2009) shows that at finite sample sizes GCV is prone to undersmoothing and is more likely to develop multiple minima than REML. Therefore, we employ REML given its practical advantages for smoothing parameter selection. We use the `gam()` function of the `R` package `mgcv` since approximate REML optimization for GAMs is reliably and efficiently implemented (Wood 2010).

Given the dimension of the dataset used here, we experienced some memory issues which could not be sorted out. This was eventually dealt with by trying to find a compromise between tensor product basis dimension and dataset reduction. The final model was fitted on data obtained by averaging over a 150×150 grid for all years, with a tensor product basis made up of a soap film with XX interior knots and a rank XX cyclic penalized cubic regression spline as the boundary curve, and a cubic regression spline with XX bases.

The spatiotemporal structure of model (2) and the lack of availability of economic variables at municipal level would suggest to fit the model by also taking into account the presence of some unexplained spatial

and autocorrelation structure in the data. In a smoothing spline context, this can be consistently achieved by using a mixed modelling approach (Breslow and Clayton 1993). This possibility was explored by using the `gamm()` function of the `mgcv` package, which iteratively calls the `lme()` function of the `nlme` package (Pinheiro et al. 2009) for maximization; no models could be fitted due to convergence failures (see, e.g., Ruppert, Wand, and Carroll (2003) and Wood (2006) for problems and limitations with this approach). To explore the sensitivity of our estimates, we tried out a fixed effect approach (Wooldridge 2002) but the results did not change significantly.

3.4 Variance estimation

Confidence intervals for the smooth components of the model can be obtained using the well known Bayesian 'confidence' intervals originally proposed by Wahba (1983) or Silverman (1985) in the univariate spline model context, and then generalized to GAMs (e.g., Gu 2002; Gu and Wahba 1993; Ruppert, Wand, and Carroll 2003). In fact, the Bayesian model representation provides a nice framework for constructing intervals, because the posterior distribution of the model parameters is known. Also, some simulation evidence suggests that these intervals have close to nominal across-the-function frequentist coverage probabilities (e.g., Wang and Wahba, 1995). This is because the intervals include both a bias and variance component (Nychka, 1988).

The large sample posterior used for interval calculations is

$$\boldsymbol{\theta}|\mathbf{y} \sim N(\hat{\boldsymbol{\theta}}, \mathbf{V}_{\boldsymbol{\theta}}),$$

where $\hat{\boldsymbol{\theta}}$ is the maximum penalized likelihood estimate of $\boldsymbol{\theta}$ which is of the form $(\mathbf{X}^{\mathsf{T}}\mathbf{W}\mathbf{X} + \mathbf{S})^{-1}\mathbf{X}^{\mathsf{T}}\mathbf{W}\mathbf{z}$, $\mathbf{V}_{\boldsymbol{\theta}} = (\mathbf{X}^{\mathsf{T}}\mathbf{W}\mathbf{X} + \mathbf{S})^{-1}\phi$, $\mathbf{X}$ contains the columns associated with the regression spline bases used to set up the model, and $\mathbf{W}$ and $\mathbf{z}$ are the diagonal weight matrix and the pseudodata vector at convergence of the algorithm used to fit the penalized model (e.g., Wood 2006, 2010).

Given the result above, it is easy to find confidence intervals for linear functions of the parameters such as the smooth components of the model. Furthermore, Bayesian confidence intervals for non-linear functions of the model coefficients can be readily obtained by simulation from the posterior distribution of $\boldsymbol{\theta}$.

4 Empirical results

LUCA will write comments about the results, to be discussed with GIAMPIERO afterwords and ultimately with DAVE.

4.1 Spatiotemporal trends

As above.

4.2 Foreigner presence and spatial attractiveness: is there a nexus?

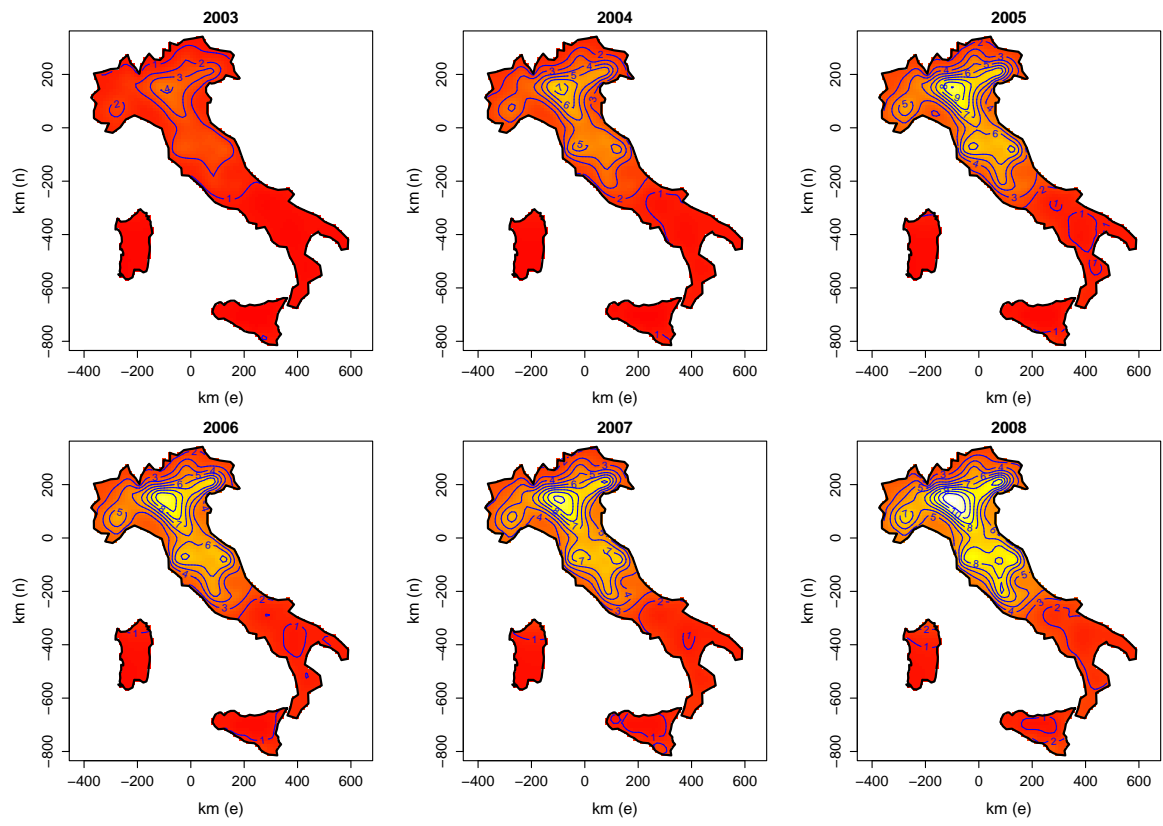As above. In addition there are some elaborations to complete.

Figure 2: to fill in

4.3 Policy-maker's corner: a point of discussion

As above.

5 Conclusions

–

References

[1] Algan, Y., Dustmann, C., Glitz, A., and Manning, A. (2010), "The economic situation of first- and second-generation immigrants in France, Germany, and the UK," *The Economic Journal*, 120, F4–F30.

[2] Bauer, T., and Zimmermann, K. F. (1999), "Assessment of possible migration pressure and its labour market impact following EU enlargement to central and eastern europe," IZA Research Report Series 3, Bonn.

[3] Borjas, J. G. (1989), "Economic theory and international migration," *International Migration Review*, 23, 457–85.

[4] Borjas, J. G. (1994), "The economics of immigration," *Journal of Economic Literature*, 32, 1667–1717.

[5] Borjas, J. G., Freeman, R. B., and Katz, L. F. (1996), "Searching for the effect of immigration on the labor market," *American Economic Review*, 86, 246–51.

[6] Borjas, J. G. (2003), "The labor demand curve IS downward sloping: reexamining the impact of immigration on the labor market," *The Quaterly Journal of Economics*, 118, 1335–74.

[7] Borjas, J. G. (2005), "The labor market impact of high-skill immigration," *American Economic Review*, 92, 56–60.

[8] Breslow, N. E., and Clayton, D. G. (1993), "Approximate inference in generalized linear mixed models," *Journal of the American Statistical Association*, 88, 9–25.

[9] Casey, T., and Dustmann, C. (2010), "Immigrants' identity, economic outcomes and the trasmission of identity across generations," *The Economic Journal*, 120, F31–F51.

[10] Card, D. (2005), "Is the new immigration really so bad?," *The Economic Journal*, 115, 300–323.

[11] Cheong, P. H., Edwards, R., Goulbourne, H., and Solomos, J. (2007), "Immigration, social cohesion and social capital: A critical review," *Critical Social Policy*, 27, 24–49.

[12] Contucci, P., and Ghirlanda, S. (2007), "Modelling society with statistical mechanics: an application to cultural contact and immigration," *Quality and Quantity*, 41, 569–78.

[13] Craven, P., and Wahba, G. (1979), "Smoothing noisy data with spline functions," *Numerische Mathematik*, 31, 377–403.

[14] De Bartolo, G. (2007), "Immigration in Italy: the great emergency," Population Association of America - Annual Meeting, March 29-31, New York, Office of Population Research, Princeton University.

[15] Edin, Per-A., Fredriksson, P., and Aslund, O. (2003), "Ethnic enclaves and the economic success of immigrants - evidence from a natural experiment," *The Quarterly Journal of Economics*, 118, 329–57.

[16] Fullin, G., and Reyneri, E. (2010), "Low unemployment and bad jobs for new immigrants in Italy," *International Migration*, In press, DOI: 10.1111/j.1468-2435.2009.00594.x.

[17] Gu, C. (2002), *Smoothing Spline ANOVA Models*, London: Springer-Verlag.

[18] Gu, C., and Wahba, G. (1993), "Smoothing Spline ANOVA with Component-Wise Bayesian Confidence Intervals," *Journal of Computational and Graphical Statistics*, 2, 97–117.

[19] Hastie, T., and Tibshirani, R. (1990), *Generalized Additive Models*, London: Chapman & Hall.

[20] Hooghe, M., Trappers, A., Meuleman, B. and Reeskens, T. (2008), "Migration to European countries: a structural explanation of patterns 1980-2004," *International Migration Review*, 42, 476–504.

[21] Lazear, E. P. (1999), "Culture and language," *Journal of Political Economy*, 107, S95–S126.

[22] Lowell, L. B. (2007), "Trends in international migration flows and stocks, 1975-2005," OECD Social, Employment and Migration Working Paper 58, OECD, Paris.

[23] Manning, A. (2010), "Feature: the integration of immigrants and their children in Europe: introduction," *The Economic Journal*, 120, F1–F3.

[24] Marra, G, and Radice, R. (2010), "Penalised regression splines: theory and application to medical research," *Statistical Methods in Medical Research*, 19, 107–125.

[25] Massey, D. S., Arago, J., Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J. E. (1993), "Theory of international migration: a review and appraisal," *Population and Development Review*, 19, 431–66.

[26] Morawska, E. (1990), "The Sociology and historiography of Immigration," in *Immigration Reconsidered: History, Sociology, and Politics*, ed. Virginia Y. McLaughlin, New York: Oxford University Press, 187–240.

[27] Nychka, D. (1988), "Bayesian Confidence Intervals for Smoothing Splines," *Journal of the American Statistical Association*, 83, 1134–1143.

[28] Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., and the R Core Team (2009), *nlme: Linear and Nonlinear Mixed Effects Models*, R package version 3.1- 96.

[29] Piore, M. J. (1979), *Birds of Passage: Migrant Labor in Industrial Societies*. Cambridge: Cambridge University Press.

[30] Portes, A., and Walton, J. (1981), *Labor, Class, and the International System*. New York: Academic Press.

[31] Ramsay, T. (2002), "Spline smoothing over difficult regions," *Journal of the Royal Statistical Society Series B*, 64, 307–19.

[32] Raphael, S., and Smolensky, E. (2009), "Immigration and poverty in the United States," *American Economic Review*, 99, 41–4.

[33] Reiss, P. T., and Ogden, R. T. (2009), "Smoothing parameter selection for a class of semiparametric linear models," *Journal of the Royal Statistical Society Series B*, 71, 505–524.

[34] Ruppert, D., Wand, M. P., and Carroll, R. J. (2003), *Semiparametric Regression*, London: Cambridge University Press.

[35] Silverman, B. W. (1985), "Some Aspects of the Spline Smoothing Approach to Non-Parametric Regression Curve Fitting," *Journal of the Royal Statistical Society Series B*, 47, 1–52.

[36] Slack, T., and Jensen, L. (2007), "Underemployment across immigrant generations," *Social Science Research*, 36, 1415–30.

[37] Stark, O., and Bloom, E.D. (1985), "The new economics of labor migration," *American Economic Review*, 75, 173–78.

[38] Wahba, G. (1983), "Bayesian 'Confidence Intervals' for the Cross-Validated Smoothing Spline," *Journal of the Royal Statistical Society Series B*, 45, 133–150.

[39] Wahba, G. (1985), "A Comparison of GCV and GML for Choosing the Smoothing Parameter in the Generalized Spline Smoothing Problem," *The Annals of Statistics*, 13, 1378–1402.

[40] Wallerstein, I. (1974), *The Modern World System, Capitalist Agricolture and the Origins of the European World Economy in the Sixteenth Century*. New York: Academic Press.

[41] Wang, Y., and Wahba, G. (1995), "Bootstrap Confidence Intervals for Smoothing Splines and Their Comparison to Bayesian Confidence Intervals," *Journal of Statistical Computation and Simulation*, 51, 263–279.

[42] Wood, S. N. (2006), *Generalized Additive Models: An Introduction with R*, London: Chapman & Hall.

[43] Wood, S. N., Bravington, M. V., and Hedley, S. L. (2008), "Soap film smoothing," *Journal of the Royal Statistical Society Series B*, 70, 931–55.

[44] Wood SN (2010), "Fast stable REML estimation of semiparametric GLMs," *Journal of the Royal Statistical Society Series B*, in press.

[45] Wooldridge, J. M. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge: MIT Press.

[46] Zanin, L., and Marra, G. (2010), "Modelling the probability that a household participates in tourism: the Italian case," *Physica A: Statistical Mechanics and its Applications*, ?.

[47] Zincone, G. (2006), "The making of policies: immigration and immigrants in Italy," *Journal of Ethnic and Migration Studies*, 32, 347–75.