Local and Global Analysis of the Fertility Rate in Italy

Massimo Mucciardi
Department of Economics
University of Messina (Italy)
massimo.mucciardi@unime.it

Mary Ellen Toffle
Department of Cognitive Science
University of Messina (Italy)
mtoffle@unime.it

1. INTRODUCTION

The Total Fertility Rate (TFR) is one of the most famous demographic indicators because it indicates the status of well-being of the population as well as an understanding of the productive potential and development of a nation.

Following the rapid and intense decrease after the "babyboom", today Italy is one of the nations with the lowest fertility rate, both at the European and world level among those nations considered to be advanced. In recent time, the TFR was equal to 1.34, very close to the so-called "lowest low fertility" threshold. These low values of fertility not only impact on the population's age structure, but also on social welfare systems since the increasing share of elderly needs more financial aid while the support from the working age population decreases.

But is this decline at the national level? Do some determinants of the TFR act more in a specific territory? Through the use of a Semi-parametric Geographically Weighted Regression (S-GWR) we try to investigate the "local and global" socio-economic determinates of TFR in Italy. The data employed in the analysis comes from a georeferenced provincial data set.

2. A BRIEF BACKGROUND OF TFR IN ITALY

As stated in the introduction, as compared to the rest of Europe, Italy is characterized by some anomalies, or delays. The low fertility and low participation of the women in the labor market are two examples. Also the percentage of women with a university education is low: against a European average of around 36%, the Italian average is at 24% (data source: Eurostat-Istat), although there is a strong recovery among the younger generations. Education, together with female employment, is an important variable for the explanation of fertility (Calamo et al, 2014; De Santis, 2011; Righi, 2003, Trimarchi, 2011). There are relatively few childcare services for children under 3 years of age (nursery school): only 54% of the total number of municipalities offer the service (Istat, 2015). Furthermore, Italy has always distinguished itself with a high marriage rate (in this research taken as a proxy of family stability) in respect to the rest of Europe (3.2 marriage per thousand inhabitants in 2015). However, it is necessary to distinguish the different regions of the country: high levels of marriage are traditionally recorded in South and Central Italy; in the North rates are lower.

In the next section we will try to model the TFR with the variables described above[1].

3. METHODOLOGY AND DATA

This section presents the methodology adopted and the data used. As we know, standard Geographically Weighted Regression (GWR) captures locally varying processes to better understand the drivers of the spatial distribution of the dependent variable. The Semi-parametric Geographically

[1] A first analysis on the spatial variation of the TFR in Italy was carried out in 2010 (see Mucciardi et al., 2010).

Weighted Regression (S-GWR), in contrast, has both geographically varying coefficients and fixed coefficients in the same model (Fotheringham et al. 2002; Nakaya et al. 2009, Nataka 2015).

The S-GWR model can be defined as follows:

$$y_i = \sum_k \beta_k (u_i, v_i) x_{k,i} + \sum_l \alpha_l z_{l,i} + \varepsilon_i \quad (1)$$

where y_i is the dependent variable, $x_{k,i}$ is the k-th independent variable with a local coefficient in the location i, (u_i, v_i) is the x-y coordinate of the location i, $z_{l,i}$ is the l-th independent variable with a fixed coefficient α_l and ε_i is the Gaussian error at location i.

The data employed in the analysis comes from geo-referenced provincial data. This geodatabase is formed by 110 Italian provinces based on administrative boundaries. The explanatory variables are taken from the Italian National Institute (ISTAT) and include, in addition to TFR (figure 1), the marriage rate (MR), female labor participation (ER_W), percentage of degree (DE) and percentage of childcare services for children under 3 years of age in the municipalities of the total municipalities of the province (NURSERY). The data was appropriately geocoded and processed using GWR 4 (estimates) and Arcgis software (maps). For reasons of data availability, the reference year for all database variables is 2015.

4. RESULTS AND CONCLUSION

In the first step we check which of the variables vary at the local level and which do not. Based on the geographical variability test (GVT) (Nakata, 2015) we set MR, DE and ER_W as local variables while MU as global or fixed variable in the S-GWR model (see GVT in the table 1). For the calibration of the S-GWR model we have considered the centroids of the provinces for the distance calculation between spatial units and the fixed

kernel technique for bandwidth selection (Fotheringham et al., 2002). Adjusted R2 values showed significant improvement in GWR model (adjusted R2 = 0.65). AICc values using the S-GWR model (AICc = -217) are much less than those in the global regression model (AICc =-181, adjusted R2 = 0.26) [2]. Moreover, GVT shows that NURSERY has a slight positive value that indicates this variable is better assumed as global[3].

Now we are going to comment on the results of the single determinants. The MR variable, employed here as proxy of family stability, is positively correlated with TFR (0.0809) but this correlation is stronger in the central-southern provinces and in the northeast of Italy. (see table 1 and figure 2). The situation of the DE variable is more complex. As reported in table 1, the OLS model shows a (non-significant) negative correlation between TFR and DE (-0.0003), but S-GWR model reveals the existence of a greater negative relationship in the south of Italy discouraging (probably) families from having children (see table 1 and figure 3). However, more social analyses should be done to explain this phenomenon. The action of the ER_W variable seems dual. Even if the OLS estimate is slightly positive (0.0047), in Northern Italy the variable is strongly positively correlated with the TFR, while in the South of Italy the traditional low values of female employment seem to discourage fertility (see table 1 and figure 4). The NURSURY variable reveals the existence of a positive significant relationship in the global model with fertility, even if the local model doesn't provide any significant spatial non-stationarity (see the small positive value in the GVT in Table 1). This result is very important and confirms the positive effect of

[2] Moreover, the Moran's I test on the residuals after fitting OLS model suggests that there is strong signal of spatial autocorrelation among the residuals (Moran's I = 0.41; p-value<0.001). So the independence assumption of the error term appears to be violated.

[3] Positive value of GVT suggests no spatial variability in terms of model selection criteria (see Nakata, 2015 for more details).

the percentage of childcare services for children under 3 years of age in the municipalities on TFR for all the Italian territory without significant territorial differences.

Table 1: Estimates and test for the OLS and S-GWR model

Variable	OLS	Min	Lower quartile	Median	Upper quartile	Max	GVT
Intercept	0.7577**	0.4181	0.9147	1.1006	1.2326	1.5971	-40
MR	0.0809**	-0.0492	-0.0294	-0.0017	0.0265	0.1240	-87
DE	-0.0003	-0.0141	-0.0053	-0.0024	-0.0011	0.0023	-27
ER_W	0.0047**	-0.0063	0.0015	0.0059	0.0088	0.0122	-92
NURSURY	0.0011**	-	-	-	-	-	0.45

Global regression results (OLS): AICc=-181; Adj-R-square=0.26; *=p<0.05; **p<0.01
S-GWR results: bandwidth= 156 Km; AICc=-217; Adj-R-square=0.65; GWR ANOVA Test: F=4.06

In conclusion, the findings of this research are generally consistent with previous studies in some relevant inquiries about TFR. Clearly this is a research "work in progress" and we hope to achieve new results. From a methodological point of view, the analysis considers "local" and "global" determinates to improve the analysis of the TFR in Italy. We have enough evidence to assert that the local model (S-GWR) performed better and provided significant improvement over the global regression model (OLS). Global statistical methods like OLS sometimes may ignore the local information and, therefore, present a false relationship (Mucciardi et al, 2010).

Fig.1: Spatial distribution of TFR by quintiles range in Italy (year 2015)

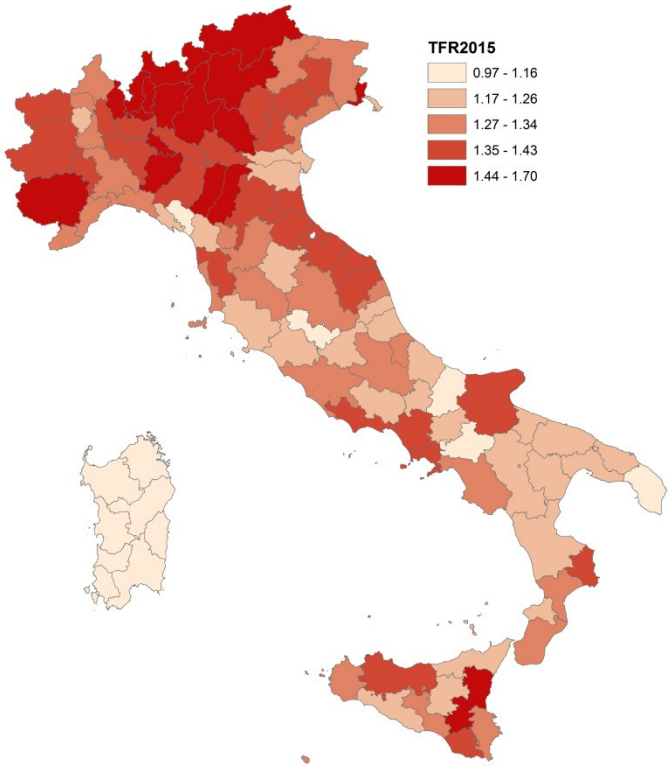


Fig.2: Local coefficient estimates of MR by quintiles range

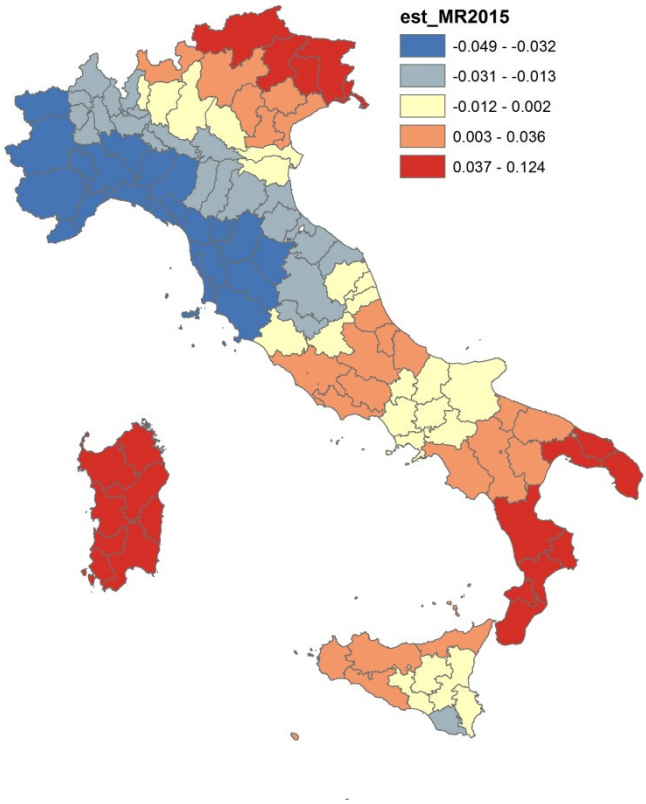


Fig.3: Local coefficient estimates of DE by quintiles range

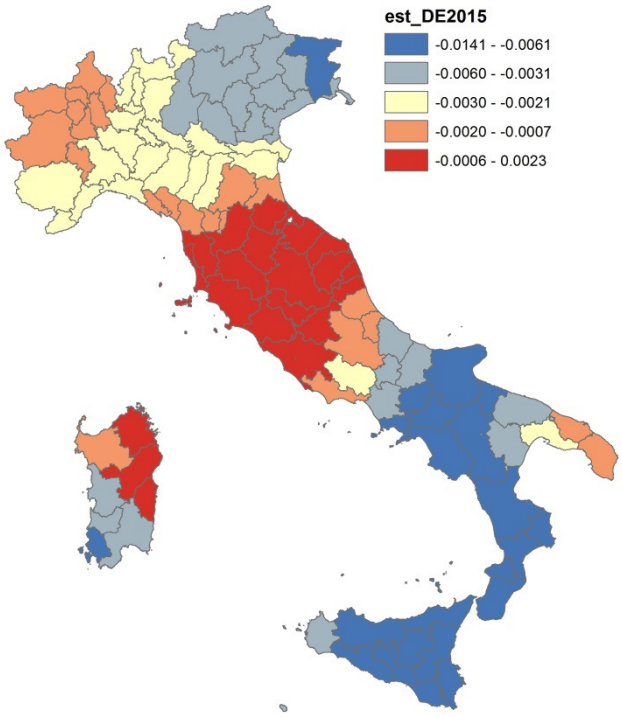
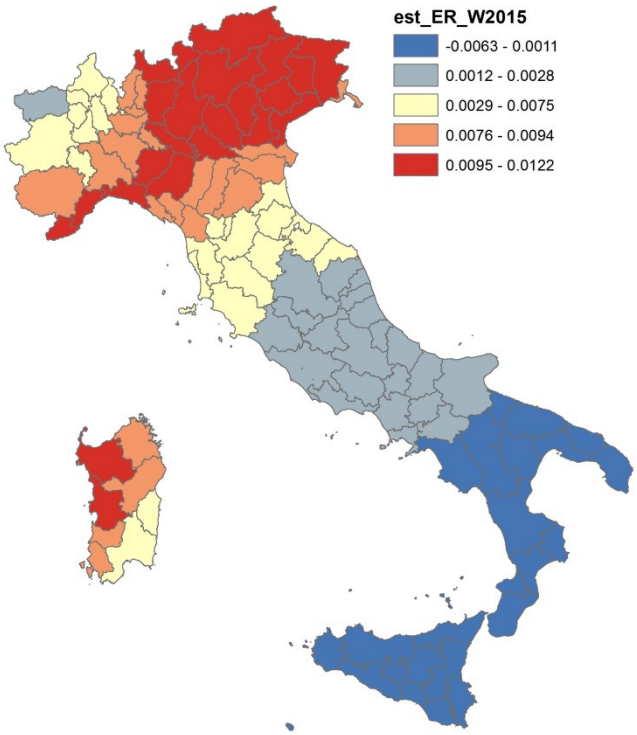


Fig.4: Local coefficient estimates of ER_W by quintiles range



REFERENCES

Calamo, R, Pereiro, T, G, (2014): RIEDS, Vol LXVIII n.1.

De Santis, G, (2011): Vi presento la mia famiglia, Neodemos http://www.neodemos.info/articoli/vi-presento-la-mia-famiglia/.

Eurostat-datasorce (2015): https://ec.europa.eu/eurostat/data/database.

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

ISTAT-datasource (2015): http://dati.istat.it/Index.aspx.

Mucciardi, M, Bertuccelli, P, (2010): Modelling Spatial Variations of Fertility Rate in Italy, Studies in Classification, Data Analysis, and Knowledge Organization, Springer.

Nakaya, T, Fotheringham, AS, Brunsdon, C, Charlton, M, (2005): Geographically weighted Poisson regression for disease associative mapping, Statistics in Medicine 24, 2695-2717.

Nakaya, T, (2015): Semiparametric geographically weighted generalised linear modelling: the concept and implementation using GWR4. Brunsdon, C and Singleton, A eds. Geocomputation: A Practical Primer, Sage Publication, 201-220.

Righi, A, (2003): Le tendenze di fecondità e di partecipazione femminile al mercato del lavoro, Roma, Istat.

Trimarchi, A, (2011): Dove più si studia si fanno più figli? Un'analisi della fecondità delle regioni italiane, Neodemos http://www.neodemos.info/articoli/dove-pi-si-studia-si-fanno-pi-figli-unanalisi-della-fecondit-delle-regioni-italiane/