REVSTAT – Statistical Journal
Volume 13, Number 1, March 2015, 1–20

ASSESSING THE EVOLUTION OF TERRITORIAL DISPARITIES IN HEALTH

Authors: Daniela Cocchi
– Department of Statistical Sciences, University of Bologna, Italy
daniela.cocchi@unibo.it

Fedele Greco
– Department of Statistical Sciences, University of Bologna, Italy
fedele.greco@unibo.it

Francesco Scalone
– Department of Statistical Sciences, University of Bologna, Italy
francesco.scalone@unibo.it

Received: October 2014 Revised: November 2014 Accepted: December 2014

Abstract:

- The paper investigates spatio-temporal trends in health disparities through an empirical example. We deal with geographical health pattern in Italy from 1991 to 2010, starting from infant mortality data available at the provincial level and assessing the existent disparity among macro-regions (the conventional Northern, Central and Southern macro-regions). After a discussion concerning suitable inequality indices and their decompositions when dealing with small area data, we propose a model-based approach that allows to properly tackle sampling variability. Results give evidences of persisting spatial disparity in infant mortality along time.

Key-Words:

- *health inequality; Gini index; generalised entropy; hierarchical Bayesian modelling; small area estimation; spatio-temporal smoothing; sampling variability; INLA.*

AMS Subject Classification:

- 62F15, 62P25, 91B72, 91B82.

1. INTRODUCTION

In this note, we investigate spatio-temporal trends in health disparities, expressed by infant mortality, through an empirical example. We deal with geographical health pattern in Italy from 1991 to 2010, starting from data available at the provincial level and assessing the existent disparity among macro-regions (the conventional Northern, Central and Southern macro-regions). The evaluation of the temporal evolution of inequality requires the adoption of suitable indicators that, along with their decomposition, help in answering a couple of main questions: is inequality between small geographical units decreasing during the study period? Which is the trend of the inequality share explained by grouping the smaller geographical units in macro-regions? As in recent years spatial disparities are being investigated in depth, the above research questions are more and more crucial. Answers to such questions are critical to improve and implement better public policies.

As a matter of fact, persistent health inequalities in modern welfare states represent a great disappointment in public health, with widening disparities reported in many Western European countries (Mackenbach, 2012). Since health inequality could persist within the same country among different regions, appropriate statistical methods to describe the spatio-temporal evolution of this phenomenon are desirable. In this study, we propose suitable decompositions of inequality indices equipped with uncertainty measures as the mean to evaluate the temporal evolution of health inequality in Italy along a twenty-years period. In particular, we focus on infant mortality, that is one of the main indicators to measure the general health level. Health disparities can be described by means of a variety of statistical measures, such as dispersions measures or inequality indices (Wagstaff *et al.*, 1991). In order to assess the presence of geographic disparities in morbidity and mortality, various authors suggested measuring health inequality by means of the Generalized Entropy and Gini indices.

In the Italian case, despite a general declining trend, some studies found high dispersion in Infant Mortality Rates (IMRs) at provincial level, revealing evident and persisting geographical disparity in infant mortality. This persisting disparity was mainly related to differences in socio-economic and health care standards among Northern, Central and Southern Italian macro-regions (Fantini *et al.*, 2005; Lauria and De Stavola, 2003).

When studying provincial-level infant mortality, IMRs show large random fluctuations giving rise to relevant methodological issues concerning the evaluation and decomposition of health disparities: more precisely, because of the low birth counts observed in the provinces and because of the rarity of infant deaths, Italian provinces have to be considered as small areas and direct estimates (*i.e.* IMRs) are subject to high sampling variability that we will tackle

by means of a model-based approach. The consequences of sampling variability on the measurement and decomposition of appropriate inequality measures constitute the main methodological focus of the paper. Literature concerning sampling variability of Gini index and Generalised Entropy measures are centred on the classical case where available data constitute a sample from a larger population, but individual values of the study variable are considered as measured without error. Instead, due to the peculiarities of the motivating example, we address the situation where the whole population has been observed (*i.e.* data concerning birth and death counts are available for each province), but individual values of the study variable need to be estimated. In this situation, the sampling properties of inequality indices estimators depend on the sampling properties of individual-level estimators. According to our knowledge, this topic has been neglected in the literature concerning health inequalities. Proper smoothing techniques need to be used in order to limit potential biases due to sample variability (Congdon *et al.*, 2001; Congdon, 2010). Adopting a popular approach to spatio-temporal disease mapping (Kim and Lim, 2010; Knorr-Held, 2000; Blangiardo *et al.*, 2013), we estimate a Bayesian smoothing model exploiting spatial association of provincial IMRs and temporal correlation. The model allows reciprocal borrowing strength for area-level data, with the least reliable rates (based on the smallest birth counts) being mostly smoothed. Model fitting is performed via Integrated Nested Laplace Approximations (INLA, Rue and Martino, 2009).

The outline of this paper is as follows. Section 2 provides a brief description of the data concerning Italian infant mortality. Section 3 introduces inequality measures and their decomposition. A simulation study is discussed for highlighting some relevant features of inequality estimators. Section 4 describes the Bayesian spatio-temporal model adopted for smoothing mortality rates: computational details are provided. In Section 5, model-based inequality decomposition is presented. In the concluding section, evidences of persisting disparity in infant mortality are briefly discussed, illustrating the contribution of the differences among macro-regions.

2. MOTIVATING EXAMPLE

Yearly data about infant mortality, in our study referred to 95 provinces along 20 years (1991–2010), are published by the Italian Institute of Statistics (ISTAT). At each year $t = 1, ..., T$, province $s = 1, ..., S$, and macro-region $k = 1, ..., K$, infant death counts y_{skt} and birth counts P_{skt} are available. For each year, province and macro-region, the Infant Mortality Rate (IMR):

$$\hat{\theta}_{skt} = \frac{y_{skt}}{P_{skt}} , \qquad s = 1, ..., S, \quad k = 1, ..., K, \quad t = 1, ..., T , \tag{2.1}$$

is a quick measure of the infant mortality intensity. In particular, $\hat{\theta}_{skt}$ is the maximum likelihood estimator of the true mortality rate θ_{skt} according to the model

$$y_{skt}|\theta_{skt} \sim Poisson(\theta_{skt}P_{skt}) \ , \qquad s = 1, ..., S, \ \ k = 1, ..., K, \ \ t = 1, ..., T \, . \tag{2.2}$$

Estimator (2.1) is unbiased and its sampling variance is inversely proportional to the birth count P_{skt}, in fact $V(\hat{\theta}_{skt}|\theta_{skt}) = \theta_{skt}/P_{skt}$. It turns out that estimates referred to provinces with low birth counts are affected by huge sampling variability, while estimates based on high birth counts are more stable. This well-known feature of mortality rates gave rise to an extensive literature concerning spatial and spatio-temporal disease mapping, aiming at smoothing observed rates (Kim and Lim, 2010; Knorr-Held, 2000; Blangiardo *et al.*, 2013). The consequences of sampling variability on the measurement and decomposition of appropriate inequality measures are discussed in the following section and constitute the main methodological focus of the paper.

Figure 1 plots the IMRs series of all Italian provinces during the study period (reported in gray). Black lines refer to IMRs observed in the Northern, Central and Southern macro-regions. A general decline of the mortality level is observed along the study period in all the macro-regions, reflecting the general trend at the provincial level. At the national level, IMR declines from 0.008 to 0.003, but a general decline in the mortality intensity does not imply a decline in territorial inequalities.

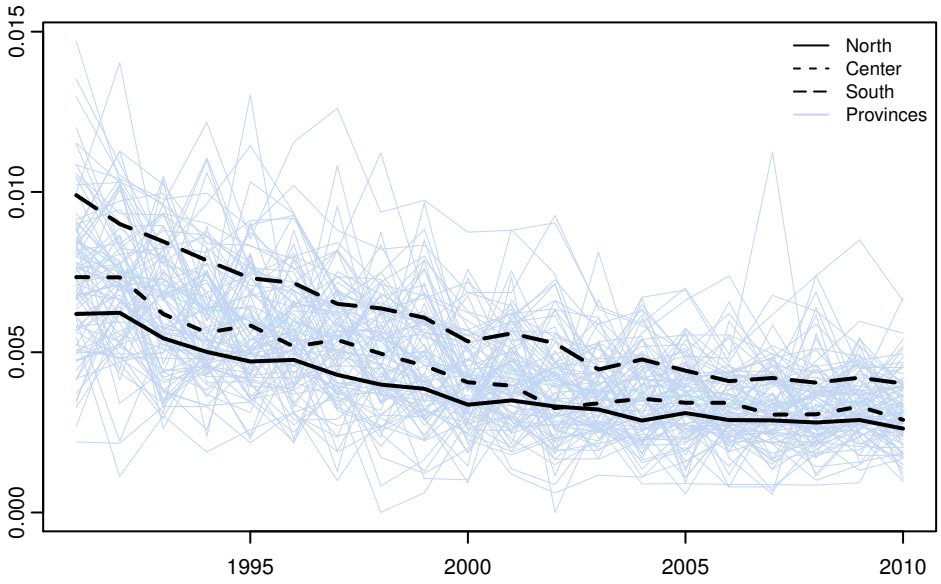

Figure 1: Temporal trend of IMRs at provincial and macro-region level.

Figure 2 reports the spatial distribution of IMRs classifying each province according to the octiles identified for the selected years 1991, 1994, 1998, 2002, 2006 and 2010. A persistent spatial trend occurs since southern provinces systematically register higher infant mortality with respect to northern provinces. In the following of this paper, we discuss how this territorial disparity can be evaluated focusing on both the overall inequality and the share of inequality explained by grouping provinces in macro-regions.

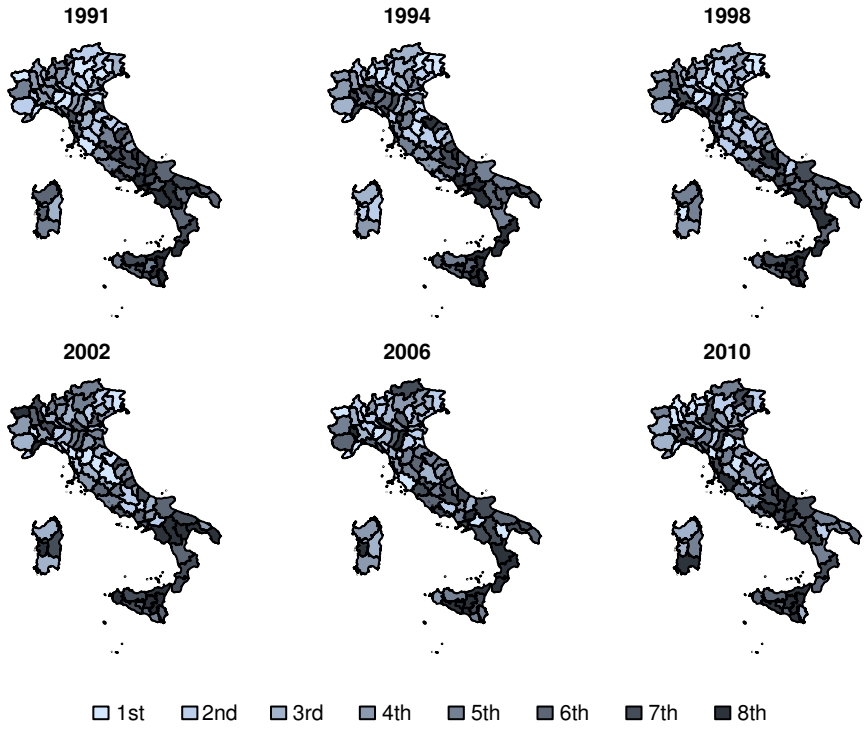

Figure 2: Octiles of the IMRs spatial distribution in selected years.

3. INEQUALITY MEASURES AND THEIR DECOMPOSITION

The theory concerning the measurement of inequalities has a long history and has been developed essentially in the framework of income distribution (Dalton, 1920; Atkinson, 1970; Dreher and Gaston, 2008). The same theory has been subsequently turned to the study of health inequalities at several spatial and temporal levels of aggregation. A popular approach defines health inequality as the uneven distribution of health across all units in a population and in population subgroups (see *e.g.*, Gakidou and King, 2002; Pradhan *et al.*, 2003). In this section, following this approach, we consider a framework where population

units are constituted by small geographical areas grouped in macro-regions and discuss some crucial statistical properties of inequality measures estimators when small area rates are involved. In particular, we focus on two popular inequality indicators: the Generalised Entropy class of indicators and the Gini coefficient. For simplifying notation, we drop the temporal subscript in what follows.

Given a population of S areas organised in K groups, the number of areas belonging to the k-th group is denoted as S_k, such that $\sum_{k=1}^{K} S_k = S$. Let $\boldsymbol{\theta}_k = (\theta_{1k}, ..., \theta_{sk}, ..., \theta_{S_k k})$ and $\mathbf{P}_k = (P_{1k}, ..., P_{sk}, ..., P_{S_k k})$ denote respectively the "true" mortality rates and the number of births referred to group k. Group-specific rates are weighted averages of area-specific rates denoted as $\bar{\theta}_k = \sum_{s=1}^{S_k} P_{sk}\theta_{sk}/P_k$, where $P_k = \sum_{s=1}^{S_k} P_{sk}$.

The Generalised Entropy is defined as:

$$GE(\boldsymbol{\theta}; \alpha) = \frac{1}{\alpha(\alpha-1)} \sum_{k=1}^{K} \sum_{s=1}^{S_k} \frac{P_{sk}}{P} \left(\left(\frac{\theta_{sk}}{\bar{\theta}} \right)^{\alpha} - 1 \right), \qquad \alpha \neq 0, 1 \ , \tag{3.1}$$

where α controls the weight assigned to the distance between mortality rates at different parts of the rates distribution: for negative/positive values of α, GE is more sensitive to changes in the lower/upper tail of the distribution. The GE class of inequality measures includes as special cases, among others, the Theil index ($\alpha = 0$) and the Coefficient of Variation ($\alpha = 2$, where the GE is equivalent to half times the squared coefficient of variation, or relative variance). The GE class of inequality measures is easily decomposable in the between and within group components. Namely, the between component is expressed as:

$$GE_B(\boldsymbol{\theta}; \alpha) = \frac{1}{\alpha(\alpha-1)} \sum_{k=1}^{K} \frac{P_k}{P} \left(\left(\frac{\bar{\theta}_k}{\bar{\theta}} \right)^{\alpha} - 1 \right),$$

a weighted average of the distances between the group means and the overall mean. The within component is expressed as a linear combination of the GEs in each sub-group

$$GE_W(\boldsymbol{\theta}; \alpha) = \sum_{k=1}^{K} \frac{P_k}{P} \left(\frac{\bar{\theta_k}}{\bar{\theta}} \right)^{\alpha} GE_{Wk} \ ,$$

where GE_{Wk} is the GE in the k-th group:

$$GE_{Wk} = GE(\boldsymbol{\theta}_k; \alpha) = \frac{1}{\alpha(\alpha-1)} \sum_{s=1}^{S_k} \frac{P_{sk}}{P_k} \left(\left(\frac{\theta_{sk}}{\bar{\theta}_k} \right)^{\alpha} - 1 \right).$$

Eventually, GE is decomposed in the between and within components as:

$$GE(\boldsymbol{\theta}; \alpha) = GE_B(\boldsymbol{\theta}; \alpha) + GE_W(\boldsymbol{\theta}; \alpha)$$

and the contribution of grouping to the global inequality can be evaluated as the ratio:

$$GE_B(\boldsymbol{\theta}; \alpha)/GE(\boldsymbol{\theta}; \alpha) \ . \tag{3.2}$$

As pointed out in Dagum (1997), the decomposition of GE-type inequality measures is essentially based on the hypotheses underlying one-way analysis of variance, neglecting differences in variances and asymmetry characterising subgroups, and delivering the between component from comparisons between group means. A nicer and more detailed picture of inequality decomposition can be obtained starting from the Gini index defined as:

$$G(\boldsymbol{\theta}) = \frac{1}{2\bar{\theta}} \sum_{h=1}^{K} \sum_{k=1}^{K} \sum_{s=1}^{S_h} \sum_{v=1}^{S_k} \frac{P_{sh}}{P} \frac{P_{vk}}{P} |\theta_{sh} - \theta_{vk}| \tag{3.3}$$

according to the proposal of Dagum (1997). This decomposition considers three components measuring respectively within-group inequality, net between-group inequality and transvariation (*i.e.* overlapping) between groups. The component due to transvariation represents one strong peculiarity of Gini's index with respect to the GE decomposition. For simplifying notation, it is assumed that the group means are ordered as $\bar{\theta}_1 \leq ... \leq \bar{\theta}_k \leq ... \leq \bar{\theta}_K$. The decomposition starts by defining the Gini index between the couple of groups h and k as:

$$G_{hk} = \frac{1}{\bar{\theta}_h + \bar{\theta}_k} \sum_{s=1}^{S_h} \sum_{v=1}^{S_k} \frac{P_{sh}}{P_h} \frac{P_{vk}}{P_k} |\theta_{sh} - \theta_{vk}| \ . \tag{3.4}$$

For $h = k$, expression (3.4) corresponds to the Gini index of the k-th group. It immediately turns out that (3.3) can be written as a function of (3.4) as:

$$G(\boldsymbol{\theta}) = \sum_{h=1}^{K} \sum_{k=1}^{K} \frac{P_h}{P} \left(\frac{P_k \bar{\theta}_k}{P\bar{\theta}} \right) G_{hk} = \sum_{h=1}^{K} \sum_{k=1}^{K} q_h r_k G_{hk} \ , \tag{3.5}$$

where $q_h = P_h/P$ is the population share of the h-th group and $r_k = (P_k\bar{\theta}_k)/(P\bar{\theta})$ can be interpreted as the share of expected death counts in the k-th group. Since $\sum_{h=1}^{K} \sum_{k=1}^{K} q_h r_k = 1$, the Gini index can be expressed as a weighted average of the between groups Gini indices G_{hk}; on the contrary, it is not possible to express GE-based decompositions as weighted averages, since the weights do not sum up to one. Coefficients G_{hk}, properly combined with weights q_h and r_k, allow to decompose the Gini index in three components. The first one is

$$G_W(\boldsymbol{\theta}) = \sum_{k=1}^{K} q_k r_k G_{kk} \ , \tag{3.6}$$

which measures the contribution of within group inequality. The following expression of the between component is due to Costa (2009):

$$G_B(\boldsymbol{\theta}) = \sum_{h=1}^{K-1} \sum_{k=h+1}^{K} \frac{r^*_{hk} - q^*_{hk}}{r^*_{hk} q^*_{kh} + r^*_{kh} q^*_{hk}} (q_h r_k + q_k r_h) \ , \tag{3.7}$$

where $r^*_{hk} = r_h/(r_h + r_k)$ and $r^*_{hk} = q_h/(q_h + q_k)$. The component due to transvariation, denoted in what follows as $G_T(\theta)$, can be obtained by difference. The

between component of the Gini index has the merit to take into account pairwise differences between individuals instead of being entirely based on comparisons among group means: for this reasons it should be preferred to GE-like indices. Eventually, the decomposition

$$G(\boldsymbol{\theta}) = G_W(\boldsymbol{\theta}) + G_B(\boldsymbol{\theta}) + G_T(\boldsymbol{\theta}) \tag{3.8}$$

is obtained.

For the purpose of this work, we consider $G(\boldsymbol{\theta})$, $G_B(\boldsymbol{\theta})$, $G_W(\boldsymbol{\theta})$, $G_T(\boldsymbol{\theta})$, $GE(\boldsymbol{\theta};\alpha)$, $GE_B(\boldsymbol{\theta};\alpha)$ and $GE_W(\boldsymbol{\theta};\alpha)$ as target population parameters to be estimated.

3.1. The use of direct estimates

Direct estimates $\hat{\theta}_{sk}$ plugged in the expression of population quantities are a popular way for estimating inequality indices. Figures 3 and 4 report estimates $\widehat{G}(\boldsymbol{\theta})$, $\widehat{G}_B(\boldsymbol{\theta})$, $\widehat{G}_W(\boldsymbol{\theta})$, $\widehat{G}_T(\boldsymbol{\theta})$, $\widehat{GE}(\boldsymbol{\theta};\alpha)$, $\widehat{GE}_B(\boldsymbol{\theta};\alpha)$ and $\widehat{GE}_W(\boldsymbol{\theta};\alpha)$ of these inequality measures for each year in the interval 1991–2010 concerning Italian infant mortality, obtained by simply plugging-in direct estimates of the mortality intensity. As an example, at each year, $\widehat{G}(\boldsymbol{\theta})$ is obtained as:

$$\widehat{G}(\boldsymbol{\theta}) = \frac{1}{2\hat{\bar{\theta}}} \sum_{h=1}^{K} \sum_{k=1}^{K} \sum_{s=1}^{S_h} \sum_{v=1}^{S_k} \frac{P_{sh}}{P} \frac{P_{vk}}{P} |\hat{\theta}_{sh} - \hat{\theta}_{vk}| \tag{3.9}$$

employing the direct estimates (2.1).

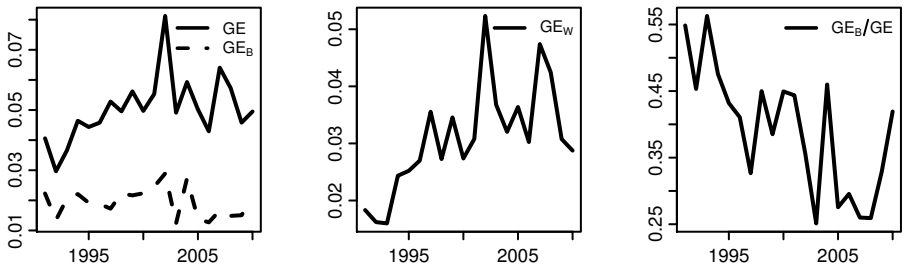

Figure 3: Plug-in estimates. Generalised Entropy and Between Generalised Entropy (left panel). Within Generalised Entropy (middle panel). Effect due to the between component (right panel).

Estimates of inequality indicators show a noisy temporal trend that suggests an increasing weight of the within components (see middle panels of the following Figures 3 and 4) and a decreasing weight of the between component

(see Figure 3 right panel and Figure 4 lower left panel). In particular, according to the GE index, the between component accounts for 55% of total inequality in 1991, decreases to 25% in year 2008 and then shows a further increase to 36% in 2010. The between component of the Gini index accounts for 70% of total inequality in 1991, decreases to 49% in year 2008 and then shows a further increase to 60% in 2010. Our purpose is to show that these estimates should not be considered as reliable pictures of territorial disparity in Italian infant mortality, since they are heavily affected by the sampling variability of direct estimates.

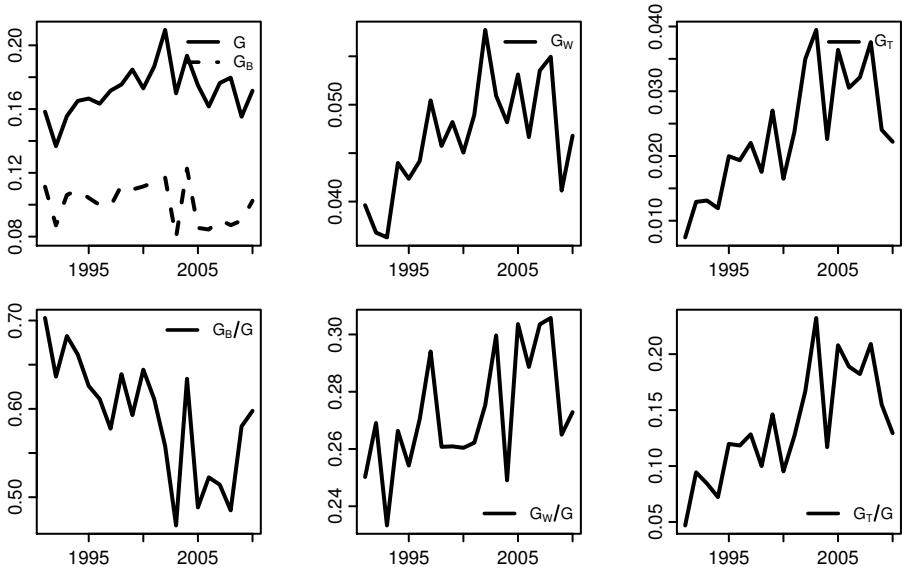

Figure 4: Plug-in estimates. Gini index and its between component (upper left panel). Within component (upper middle panel). Transvariation component (upper right panel). Lower panels report the contribution of each component to the total.

Literature concerning sampling variability of Gini index and Generalised Entropy measures focuses on the classical case where available data constitute a sample from a larger population, but individual values of the study variable are considered as measured without error (see for example Langel and Tillé, 2013). Instead, due to the peculiarities of the motivating example, we address the situation where the whole population has been observed (*i.e.* data concerning birth and death counts are available for each province), but the individual values of study variable (*i.e.* mortality intensity θ_{sk}) need to be estimated. In this situation, the sampling properties of inequality indices estimators depend on the sampling properties of individual-level estimators. According to our knowledge, this topic has been neglected in the literature concerning health inequalities. In Subsection 3.2 the effect of estimating inequality measures by simply plugging-in direct estimates $\hat{\theta}_{sk}$ is discussed.

3.2. The effect of sampling variability on decomposition

In order to discuss the consequences of direct estimates sampling variability on the estimation of inequality measures decomposition, we design a simulation study that considers a population partitioned in $K = 3$ groups, where the whole inequality is explained by the between-group component, while equality within groups is postulated. We set $\bar{\theta}_1 = 0.6\ \bar{\theta}$, $\bar{\theta}_2 = \bar{\theta}$ and $\bar{\theta}_3 = 1.4\ \bar{\theta}$. Moreover, we set $\theta_{sk} = \bar{\theta}_k\ \forall s, k$, such that the within component of any inequality measure equals 0, *i.e.* $G_W(\boldsymbol{\theta}) = GE_W(\boldsymbol{\theta}; \alpha) = 0$. In order to investigate the effect of mortality intensity, we let $\bar{\theta}$ vary between .002 and .009, similarly to the national mortality levels observed between 1991 and 2010 in Italian infant mortality. For the sake of simplicity, from now on, we set $\alpha = 1.5$ when dealing with the GE index. With this setting, for any $\bar{\theta}$, we obtain $G(\boldsymbol{\theta}) = G_B(\boldsymbol{\theta}) = 0.196$ and, fixing $\alpha = 1.5$, $GE(\boldsymbol{\theta}; 1.5) = GE_B(\boldsymbol{\theta}; 1.5) = 0.069$. For each value of $\bar{\theta}$, and for each s and k, $M = 50,000$ death counts $\{y_{sk}^m\}_{m=1,...,M}$ are generated from the model:

$$y_{sk}|\bar{\theta}_k \sim Poisson(\bar{\theta}_k P_{sk})\ , \qquad k = 1, ..., K, \quad s = 1, ..., S_k\ ,$$

where P_{sk} is set at the number of births observed in the Italian provinces in 2010, in order to obtain simulation results relevant for highlighting the peculiarities of our case-study. For each simulated count y_{sk}^m, direct estimates $\hat{\theta}_{sk}^m = y_{sk}^m / P_{sk}$ are used to obtain plug-in estimates of the inequality measures and their components. Averaging over all simulated values, we obtain the expected value of the plug-in estimators, as an example:

$$E(\widehat{G}(\boldsymbol{\theta})|\boldsymbol{\theta}) = \frac{1}{M} \sum_{i=m}^{M} \widehat{G}^m(\boldsymbol{\theta})\ .$$

Simulation results are reported in Figures 5 and 6, with $\bar{\theta}$ values in abscissa. In all panels, true population values are reported as horizontal thin lines. The left panel of Figure 5 and the upper left panel of Figure 6 show that estimators $\widehat{G}(\boldsymbol{\theta})$ and $\widehat{GE}(\boldsymbol{\theta}; \alpha)$ of the global inequality are positively biased while both estimators of the between components, whose expected value is reported as a dashed line in the same panels, are approximately unbiased. Unbiasedness of the between-component estimators is not surprising and can be ascribed to the stability of the group-specific sample means $\hat{\bar{\theta}}_k$ as estimators of the population parameters $\bar{\theta}_k$, based on greater population sizes with respect to area-level estimates $\hat{\theta}_{sk}$. It turns out that the bias of the global estimators is essentially due to overestimation of the within component, as can be seen from the central panels of the figures. Moreover, the bias of the global measures decreases when $\bar{\theta}$ increases: the relative bias of $\widehat{GE}(\boldsymbol{\theta}; 1.5)$ ranges from 63% (when $\bar{\theta} = .002$) to 14% (when $\bar{\theta} = .009$), while the relative bias of $\widehat{G}(\boldsymbol{\theta})$ ranges from 34% (when $\bar{\theta} = .002$) to 13% (when $\bar{\theta} = .009$).

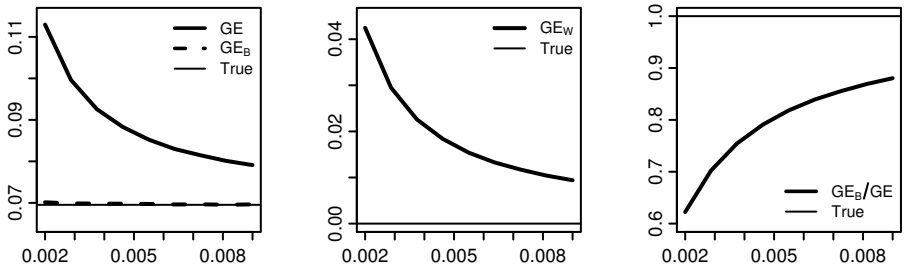

Figure 5: Expected values of the components of the Generalised Entropy index (left and central panels). Expected values of the contribution of the between component to the total (right panel). In abscissa $\bar{\theta}$ values.

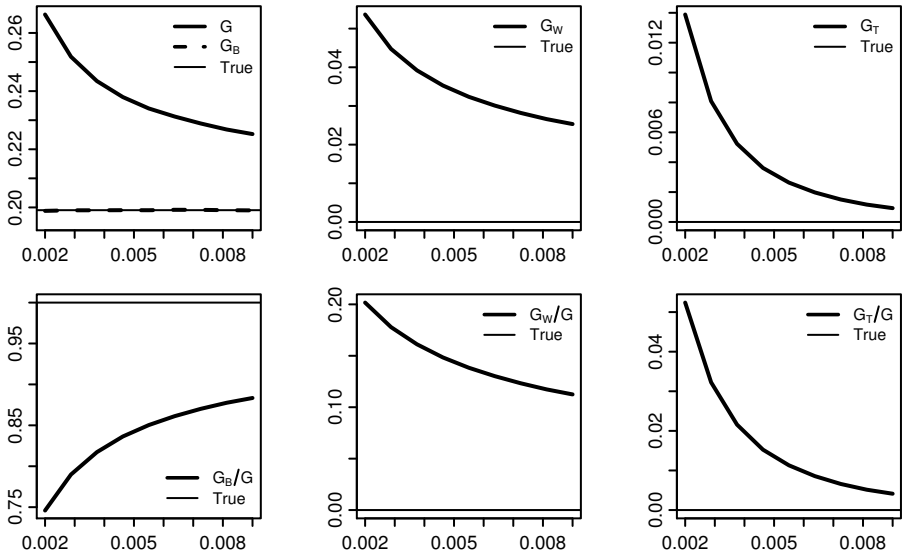

Figure 6: Expected values of the components of the Gini index (upper panels). Expected values of the contribution of each component to the total (lower panels). In abscissa $\bar{\theta}$ values.

This is a very relevant feature to bear in mind in our case study: since the average mortality intensity decreases along the study period (see Figure 1), it is very likely that overestimation of the within component is more severe at the end of the study period itself. In other words, inequality measures computed at the beginning and at the end of the period (reported in Figures 3 and 4) are not directly comparable since they are affected in a different way by sampling variability. An interesting feature of the Dagum's decomposition of the Gini index is its ability to capture (and to be affected by) transvariation: for low $\bar{\theta}$ values, simulated rates $\widehat{\theta}^m_{sk}$ are more likely overlapping between groups than for high $\bar{\theta}$ values: this intuitive behaviour induces the trend of $E(\widehat{G}_T(\boldsymbol{\theta})|\boldsymbol{\theta})$ plotted in the right panels of Figure 6. As a consequence of the overestimation of within

variability, the contribution of the between components to the global inequality turns out to be heavily underestimated (see Figure 5, right panel and Figure 6, lower left panel): $E(\widehat{GE}_B(\boldsymbol{\theta}; 1.5)/\widehat{GE}(\boldsymbol{\theta}; 1.5)|\boldsymbol{\theta})$ ranges from 0.62 to 0.88 as a function of $\bar{\theta}$, while $E(\widehat{G}_B(\boldsymbol{\theta})/\widehat{G}(\boldsymbol{\theta})|\boldsymbol{\theta})$ ranges from 0.75 to 0.89.

The dangers of a quick exploitation of direct estimates have been highlighted by the simulation study just described. A model-based approach where small area estimates are improved by a borrowing strength process is therefore needed. The Bayesian framework is particularly suitable to this aim and to easily obtain uncertainty measures concerning inequality decomposition.

4. SPATIO-TEMPORAL SMOOTHING

Spatio-temporal disease mapping models can be adopted as useful tools for attenuating the effects of sampling variability of individual-level estimates on inequality measures and their decomposition. Several spatio-temporal disease mapping models have been proposed including parametric or non-parametric time trend and different types of spatio-temporal interaction (see *e.g.* Blangiardo *et al.*, 2013; Schrödle and Held, 2011; Ugarte *et al.*, 2014). In this work, we adopt the well-known smoothing model proposed in Knorr-Held (2000), that is briefly sketched in what follows. Since our aim is limited to obtain smoothed mortality rates, we do not include group-specific parameters: between-group variation will be evaluated on the basis of the posterior distribution of the smoothed rates. According to the approach proposed in Knorr-Held (2000), the spatio-temporal trend is non-parametrically modelled: this delivers a very flexible model that can capture complex non-linear behaviours. Smoothing is achieved by borrowing strength along both space and time under the fairly reasonable hypothesis that rates variation is smooth along these dimensions. The model is hierarchically specified and is particularly suitable to be managed in a Bayesian framework. At the first level of the hierarchy, conditionally on model parameters involved in higher levels, mortality counts y_{skt} are assumed to follow independent Poisson distributions:

$$y_{skt}|\theta_{skt} \sim Poisson(\theta_{skt}P_{skt}), \qquad s = 1, ..., S, \;\; k = 1, ..., K, \;\; t = 1, ..., T\,. \tag{4.1}$$

In its most general formulation, the model includes both spatial and temporal structured and unstructured random effects and a spatio-temporal interaction term. All random effects are modelled as Gaussian Markov Random Fields (GMRF): the Markov property of GMRF models implies sparseness of the precision matrix, which allows fast computations. The linear predictor is specified as:

$$log(\theta_{skt}) = \mu + \phi_t + \nu_t + \psi_{sk} + u_{sk} + \delta_{skt}, \tag{4.2}$$

where μ captures the average log-rate; $\boldsymbol{\nu} = (\nu_1, ..., \nu_t, ..., \nu_T)$ and $\boldsymbol{u} = (u_{11}, ..., u_{S_1 1}, ..., u_{1K}, ..., u_{S_K K})$ are unstructured temporal and spatial random effects distributed as independent zero-mean Gaussian variables, *i.e.* $\boldsymbol{u} \sim N(\mathbf{0}, \tau_u \mathbf{I}_S)$ and $\boldsymbol{\nu} \sim N(\mathbf{0}, \tau_\nu \mathbf{I}_T)$. Intrinsic GMRF (IGMRF) are adopted for random effects $\boldsymbol{\phi} = (\phi_1, ..., \phi_t, ..., \phi_T)$ and $\boldsymbol{\psi} = (\psi_{11}, ..., \psi_{S_1 1}, ..., \psi_{1K}, ..., \psi_{S_K K})$, namely $\boldsymbol{\phi} \sim N(\mathbf{0}, \tau_\phi \mathbf{K}_T(\phi))$ and $\boldsymbol{\psi} \sim N(\mathbf{0}, \tau_\psi \mathbf{K}_S(\psi))$, where $\mathbf{K}_T(\phi))$ is structured in order to obtain a Random Walk 1 prior and $\mathbf{K}_S(\psi))$ depends on the neighbouring structure of the map, delivering the well-known Intrinsic Conditional AutoRegressive (ICAR) model. With regard to the spatio-temporal interaction random effects $\boldsymbol{\delta} = (\delta_{111}, ..., \delta_{skt}, ..., \delta_{SKT})$, four types of interaction can be postulated by specifying the structure matrix as the Kronecker product of the corresponding structure matrices of the main effects. Namely, $\boldsymbol{\delta} \sim N(\mathbf{0}, \tau_\delta \mathbf{K}_{ST}(\delta))$ where:

- Type I interaction: $\mathbf{K}_{ST}(\delta) = \mathbf{I}_T \otimes \mathbf{I}_S$;
- Type II interaction: $\mathbf{K}_{ST}(\delta) = \mathbf{I}_T \otimes \mathbf{K}_S(\psi)$;
- Type III interaction: $\mathbf{K}_{ST}(\delta) = \mathbf{K}_T(\phi) \otimes \mathbf{I}_S$;
- Type IV interaction: $\mathbf{K}_{ST}(\delta) = \mathbf{K}_T(\phi) \otimes \mathbf{K}_S(\psi)$.

To ensure model identifiability, appropriate linear constraints are needed for the random effects: with regard to IGMRFs, the number of required linear constraints equals the rank-deficiency of the precision matrix. As pointed out in Schrödle and Held (2011), identifiability can be ensured by computing the null space of the structure matrices and using the obtained eigenvectors as linear constraints: this is the strategy we adopt for model estimation. Unstructured random effects are constrained to zero sum in order to allow identification of the intercept term μ. Model hierarchy is completed by specifying a diffuse Gaussian distribution as a prior μ, while Gamma priors are specified for precision parameters $\tau_\phi, \tau_u, \tau_\nu, \tau_\psi$ and τ_δ.

4.1. Computations

Coherently with the Bayesian framework, we aim at evaluating and decomposing inequality measures (3.1) and (3.3) on the basis of their posterior distribution $p(G(\boldsymbol{\theta}_t)|\mathbf{y})$ and $p(GE(\boldsymbol{\theta}_t; \alpha)|\mathbf{y})$: this allows to easily obtain both point estimates and their associated uncertainty. When dealing with complex hierarchical Bayesian models, the joint posterior distribution is not available in closed form and needs to be approximated. Two alternative strategies are currently very popular for approximating the joint posterior distribution: Markov Chain Monte Carlo (MCMC) sampling and INLA (see Rue and Martino, 2009). The latter is particularly suitable for latent GMRF models and provides very accurate approximations of the posterior distribution. Moreover, INLA outper-

forms MCMC approaches in terms of computational time and accuracy. INLA has been made easily implementable by the R package INLA (Rue *et al.*, 2013), that we used for model estimation.

It is worth noting that inequality measures are non-linear combinations of model parameters: the R package INLA allows to approximate the posterior distribution of linear combinations of the model parameters, but does not allow to obtain approximations of non-linear combinations: as a consequence, posterior inference concerning inequality measures can only be performed by sampling from the *joint* posterior distribution, a task that is naturally addressed in an MCMC framework. Fortunately, the adoption of an MCMC algorithm can be avoided in our case study, since an experimental function implemented in the INLA package, `inla.posterior.sample`, allows to draw samples from the joint posterior distribution. We checked the coherence between the results obtained by the INLA experimental function and the posterior samples obtained by means of an MCMC algorithm, finding agreement between results for the estimated models, with an impressively lower computational time demanded by the INLA-based procedure. Once posterior samples from the joint posterior distribution are available, as is the case where MCMC sampling is performed, posterior distributions of any functions of the model parameters can be obtained on the basis of these samples. Given L samples $\{\boldsymbol{\theta}_t^l\}_{l=1,...,L}$ from the joint posterior distribution $\boldsymbol{\theta}_t|\mathbf{y}$, $l = 1, ..., l, ..., L$, $t = 1, ..., t, ..., T$, for each l, inequality measures and their decompositions can be computed, delivering an L-dimensional sample from their posterior distribution: as a byproduct, both posterior point estimates and credibility intervals can be easily obtained.

5. RESULTS

Model selection is performed by means of the Deviance Information Criterion (DIC, Spiegelhalter el at., 2002) according to the results of Table 1. Models without unstructured terms are preferred in terms of fitting, as the first column of results shows; the selected model includes a Type II interaction term.

Table 1: Model comparison: Deviance Information Criterion.

Interaction	**Without ν and u**	**With ν and u**
Type I	11185.09	11224.34
Type II	**11136.75**	11196.59
Type III	11245.67	11356.31
Type IV	11236.46	11286.92

On the basis of the selected model, we obtain posterior estimates of inequality measures and their decomposition of Figures 7 and 8, where posterior means are reported along with 90% credibility intervals.

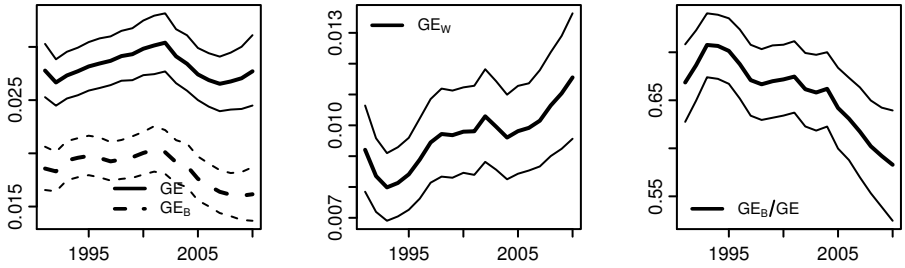

Figure 7: Posterior means and credibility intervals. Generalised Entropy and Between Generalised Entropy (left panel). Within Generalised Entropy (middle panel). Effect due to the between component (right panel).

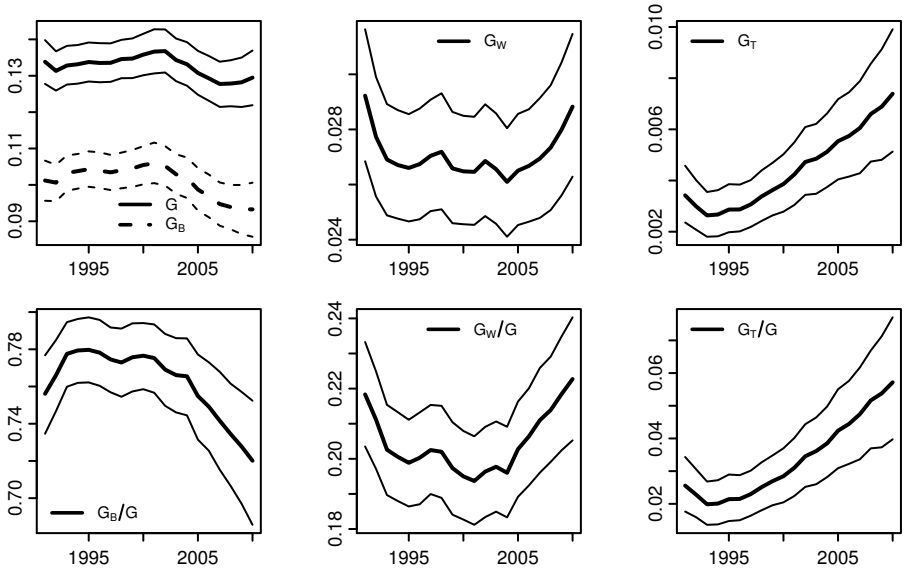

Figure 8: Posterior means and credibility intervals. Gini index and its between component (upper left panel). Within component (upper middle panel). Transvariation component (upper right panel). Lower panels report the contribution of each component to the total.

These figures should be compared with their counterparts based on direct estimates, already reported in Figures 3 and 4; all the comparisons discussed in what follows are coherent with the results of the simulation study reported in Figures 5 and 6, and should be interpreted in light of them. As a first difference,

smoothing of direct estimates turns out in smoothing of the temporal trend of inequality measures, delivering a clearer picture of the evolution of territorial disparities in Italian infant mortality. Secondly, posterior estimates of the overall level of inequality (see Figure 7 left panel and Figure 8 upper left panel) are sensibly lower than estimates obtained by means of plug-in estimators. The ratio $E(GE(\boldsymbol{\theta}_t)|\mathbf{y})/\widehat{GE}(\boldsymbol{\theta}_t)$, that can be interpreted as a quick measure of the effect due to the shrinkage of mortality rates estimates, ranges from 0.9 at the beginning of the study period, when mortality intensity is higher, to 0.45 at the end of the study period, characterised by lower mortality intensity. The same ratio referred to the Gini index ranges from about 0.9 at the beginning of the study period to about 0.7 in last years, witnessing a lower sensitivity of the Gini index to sampling variability of direct estimates. The difference between plug-in estimates and model-based posterior estimates is almost entirely due to the reduction of the within components (central panels of Figures 7 and 8) and, with regard to the Gini decomposition, to the reduction of the component measuring transvariation (Figure 8, right panel). Estimates of the between components remain basically unchanged for both indicators: as a result, the contribution of the between group variability is higher when considering model-based estimates. Despite some evidence of a decreasing trend, inequality between macro-regions explains a considerable share of global inequality: according to the Gini decomposition, which better captured the between-group component in the simulation study, such share ranges from 76% in 1991 to 72% in 2010. The same shares are reduced respectively to 66% and 57% when considering the Generalised Entropy decomposition.

6. CONCLUSIONS

In this paper, we studied the time trend of health disparity in Italy adopting a small area geographical scale. The analysis ranged over a number of methodological and empirical issues that emerge when combining traditional inequality indices, methods for their decomposition and Bayesian hierarchical models. We assumed provinces as units of analysis by grouping them in three main macro-regions: Northern, Central and Southern areas of Italy.

In order to evaluate the temporal evolution of health inequality in Italy, we focused on IMRs, since they are commonly considered as good proxies of health, environmental and socio-economic conditions. After defining, for the sake of brevity, health inequality as the uneven distribution of health across all units of a population, we took into account two popular classes of inequality indicators such as the Generalized Entropy and the Gini coefficient. We also measured the share of global inequality due to disparities among macro-regions, decomposing the total index in its basic components related to the within- and between-group inequality.

However, we preliminary showed that, when dealing with small area data, inequality measures based on direct IMRs tend to be severely affected by random fluctuations. In order to reduce the effect of sample variability and smooth direct IMRs, we estimated a Bayesian model that takes into account spatial, temporal and spatio-temporal interaction effects. Bayesian inference was carried out by means of INLA. Inequality measures based on posterior estimates come out to be less affected by random variations. The model-based Generalized Entropy and Gini coefficient appear stable over the study period, revealing a persistent inequality in infant mortality. In addition, it also comes out that the proportion of global inequality due to disparities among macro-regions tends to be higher when model based estimates are taken into account. We concluded that the persistent health disparity at provincial level is not due to small areas random variability, but is more evidently connected to relevant differences among macro-regions.

Since neonatal care given to mothers and newborns represents one of the main infant mortality causes (Scioscia *et al.*, 2007; Parazzini *et al.*, 1992), it is possible to ascribe the observed infant mortality disparity to different levels of health services (Bonati and Campi, 2005; Mazzucco *et al.*, 2011). In these terms, the persistent disparity in infant mortality between provinces may reflect the long-term socioeconomic inequalities between Northern and Southern Italy (Golini, 2014).

ACKNOWLEDGMENTS

We thank Andrea Riebler for her very useful comments and hints on the use of INLA. The research activity of Fedele Greco was funded by a FIRB 2012 grant (project no. RBFR12URQJ; title: Statistical modeling of environmental phenomena: pollution, meteorology, health and their interactions) for research projects by the Italian Ministry of Education, Universities and Research.

REFERENCES

[1] Atkinson, A.B. (1970). On the measurement of inequality, *Journal of Economic Theory*, **2**, 244–263.

[2] Bonati, M. and Campi, R. (2005). What can we do to improve child health in Southern Italy?, *PLOS Medicine*, **2**, E250.

[3] Blangiardo, M.; Cameletti M.; Baio, G. and Rue, H. (2013). Spatial and spatio-temporal models with R-INLA, *Spatial and Spatio-Temporal Epidemiology*, **7**, 39–35.

[4] Congdon, P. (2010). Modelling trends and inequality in small area mortality, *Journal of Applied Statistics*, **31**, 6, 603–622.

[5] Congdon, P.; Campos, R.M.; Curtis, S.E.; Southall, H.R.; Gregory, I.N. and Jones, I.R. (2001). Quantifying and explaining changes in geographical inequality of infant mortality in England and Wales since the 1890s, *International Journal of Population Geography*, **7**, 35–31.

[6] Costa, M. (2009). Transvariation and inequality between subpopulations in the Dagum's Gini index decomposition, *Metron*, **3**, 229–241.

[7] Dagum, C. (1997). A new approach to the decomposition of the Gini income inequality ratio, *Empirical Economics*, **22**, 515–531.

[8] Dalton, H. (1920). Measurement of the inequality of incomes, *Economic Journal*, **30**, 9, 348–361.

[9] Dreher, A. and Gaston, N. (2008). Has globalization increased inequality?, *Review of International Economics*, **16**, 3, 516–536.

[10] Fantini, M.P.; Stivanello, E.; Dallolio, L.; Loghi, M. and Savoia, E. (2005). Persistent geographical disparities in infant mortality rates in Italy (1999-2001): comparison with France, England, Germany, and Portugal, *European Journal of Health*, **16**, 4, 429–432.

[11] Gakidou, E. and Gary, K. (2002). Measuring total health inequality: adding individual variation to group-level differences, *International Journal for Equity and Health*, **1**:3.

[12] Golini, A. (2014). *Demographic crisis and economic crisis for Italian Mezzogiorno: an iceberg detached from the continent?*. President's Invited Talk at the 47th Scientific Meeting of the Italian Statistical Society, Cagliari, June, 11–13.

[13] Kim, H. and Lim., H. (2010). Comparison of Bayesian spatio-temporal models for chronic diseases, *Journal of Data Science*, **8**, 2, 189–211.

[14] Knorr-Held, L. (2000). Bayesian modelling of inseparable space-time variation in disease risk, *Statistics in Medicine*, **19**, 2555–2567.

[15] Langel, M. and Tillé, Y. (2013). Variance estimation of the Gini index: revisiting a result several times published, *Journal of the Royal Statistical Society, Series A*, **176**, 2, 521–540.

[16] Lauria, L. and De Stavola, B.L. (2003). A district-based analyses of stillbirth and infant mortality rates in Italy: 1989-93, *Paediatric Perinatal Epidemiology*, **17**, 1, 22–23.

[17] Mackenbach, J.P. (2012). The persistence of health inequalities in modern welfare states: the explanation of a paradox, *Social Science & Medicine*, **75**, 4, 761–769.

[18] Mazzucco, W.; Cusimano, R.; Macaluso, M.; La Scola, C.; Fiumanó, G.; Scondotto, S.; Cernigliaro, A.; Corsello, G.; La Torre, G. and Vitale, F. (2011). A retrospective follow up study on maternal age and infant mortality in two Sicilian districts, *BMC Public Health*, **11**:817.

[19] Parazzini, F.; Pirotta, N.; La Vecchia, C.; Bocciolone, L. and Fedele, L. (1992). Determinants of perinatal and infant mortality in Italy, *Revue D'Epidemiologie et de Santé Publique*, **40**, 1, 15–24.

[20] Pradhan, M.; Sahn D.E. and Younger, S.D. (2003). Decomposing world health inequality, *Journal of Health Economics*, **40**, 1, 15–24.

[21] Rue, H. and Martino, S. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *Journal of the Royal Statistical Society, Series B*, **71**, 2, 319–392.

[22] Rue, H.; Martino, S.; Lindgren, F.; Simpson, D. and Riebler, A. (2013). INLA: Functions which allow to perform full Bayesian analysis of latent Gaussian models using Integrated Nested Laplace Approximation, *R Package Version* 001383402327.

[23] Schrödle, B. and Held, L. (2011). Spatio-temporal disease mapping using INLA, *Environmetrics*, **22**, 6, 725–734.

[24] Scioscia, M.; Vimercati, A.; Maiorano, A.; Depalo, R. and Selvaggi, L. (2007). A critical analysis on Italian perinatal mortality in a 50-year span, *European Journal of Obstetrics & Gynecology and Reproductive Biology*, **130**, 1, 60–65.

[25] Ugarte, M.D.; Adin, A.; Goicoa, T. and Militino, A.F. (2014). On fitting spatio-temporal disease mapping models using approximate Bayesian inference, *Statistical Methods in Medical Research*, **23**, 6, 507–30.

[26] Wagstaff, A.; van Doorslaer, E. and Paci, P. (1991). Horizontal equity in the delivery of health care, *Journal of Health Economics*, **10**, 2, 251–256.