S.A.M.P.L.E.

Small Area Methods for
Poverty and Living condition Estimates

Dissemination

WP2 Small area estimation of poverty and inequality indicators

Universita di Pisa
University of Manchester
Universidad Carlos III de Madrid
Universidad Miguel Hernández de Elche

SEVENTH FRAMEWORK PROGRAMME

EU – FP7 - SSH-2007-1
Grant Agreement no 217565

SAMPLE WP2

Small area estimation of poverty and inequality indicators

By Domingo Morales

Based on the contributions of the partners

UNIPI-DSMAE: University of Pisa, Italy
CCSR: University of Manchester, United Kingdom
UC3M: University Carlos III de Madrid, Spain
UMH: University Miguel Hernández de Elche, Spain

Índice

Description of WP2

The partners

1: Universita di Pisa
3: University of Manchester
4: Universidad Carlos III de Madrid
5: Universidad Miguel Hernández de Elche

The Tasks

2.1: Estimate the c.d.f of income at small area level (1, 3)
2.2: Small area estimates of poverty with spatial models (1, 3, 4)
2.3: SAE of poverty with temporal models (5)
2.4: SAE of poverty with spatio-temporal models (1, 4, 5)

Description of WP2

2.1: Estimation of c.d.f of income at small area level

- WP2 investigates method to estimate the Cumulative Distribution Function of Income (CDFI) in each unplanned domain (total disposable household income, equivalised total disposable income - EU definition of income and modified OECD scale).

- WP2 intends to utilize M-quantile models for small area estimation.

- WP2 performs the estimation of the cumulative distribution function of the variable of interest by combining both M-quantile and random effects models with appropriate model unbiased and design consistent estimators of the distribution function.

Description of WP2

2.1-2.4: Small area estimates of poverty indicators

2.1. WP2 proposes new methodologies for estimating poverty and inequality indicators along with their accuracy measures in small areas.

(2.1a) WP2 develops small area estimates of poverty indicators that take into account the spatial correlation between neighbour areas.

(2.1b) WP2 develops small area estimates of poverty indicators using M-quantile Geographically Weighted Regression model.

2.2. WP2 develops small area estimates using data from different periods through models that "borrow strength from time".

2.3. WP2 develops small area estimates through spatial-temporal models, which "borrow strength from space and time"

Poverty measures

- Let E_{dj} be a quantitative measure of welfare for unit j in area d.

- For example $E_{dj} = R_{dj}/H_{dj}$, where
R_{dj} = total net monetary income of household j and area d,
H_{dj} = total number of normalized members of household j and area d,

$$H_{dj} = 1 + 0.5(H_{dj\geq 14} - 1) + 0.3H_{dj<14},$$

$H_{dj\geq 14}$ is the number of members aged 14 or more in (d, j),
$H_{dj<14}$ is the number of members aged 13 or less in (d, j).

- Let z be the poverty line; that is, the threshold for E_{dj} under which a person is considered as "under poverty".

- The family of poverty measures of Foster, Greer and Thorbecke (1984), called FGT poverty measures, for a small area d is

$$F_{\alpha d} = \frac{1}{N_d}\sum_{j=1}^{N_d}\left(\frac{z - E_{dj}}{z}\right)^{\alpha} I(E_{dj} < z), \quad \alpha = 0, 1, 2, \quad d = 1, \dots, D,$$

Poverty measures

- Note that

$\hookrightarrow$ $I(E_{dj} < z) = 1$ if $E_{dj} < z$ (person under poverty)
$\hookrightarrow$ $I(E_{dj} < z) = 0$ if $E_{dj} \geq z$ (person not under poverty).

- For $\alpha = 0$ we get the proportion of individuals under poverty in small area d, also called poverty incidence or head count ratio.

- The measure for $\alpha = 1$ is called poverty gap, and measures the small area mean of the relative distance to non-poverty (the poverty gap) of each individual.

- For $\alpha = 2$ the measure is called poverty severity.

Direct estimators of poverty measures

- The direct estimators of the FGT measures are

$$f_{\alpha d}^{w}=\frac{1}{\hat{N}_d}\sum_{j\in s_d} w_{dj}\left(\frac{z-E_{dj}}{z}\right)^{\alpha} I(E_{dj}<z),\quad \alpha=0,1,2,\quad d=1,\ldots,D,$$

where

$\hookrightarrow$ $\hat{N}_d=\sum_{j\in s_d} w_{dj}$ is the direct estimator of the population size N_d of small area d.

$\hookrightarrow$ w_{dj} is the sampling weight (inverse of the probability of inclusion) of individual j in the sample from small area d

EB method for poverty estimation

- **Assumption:** There exists a transformation $Y_{dj} = T(E_{dj})$ of the welfare variables E_{dj} which follows a normal distribution.

- Poverty measure as a function of transformed variables:

$$F_{\alpha d} = \frac{1}{N_d} \sum_{j=1}^{N_d} \left\{ \frac{z - T^{-1}(Y_{dj})}{z} \right\}^{\alpha} I\left\{ T^{-1}(Y_{dj}) < z \right\} = h_\alpha(\boldsymbol{y}_d),$$

where $\boldsymbol{y}_d = (Y_{d1}, \ldots, Y_{dN_d})'$.

- **Best estimator:** The estimator of $F_{\alpha d}$ with minimum MSE is

$$\hat{F}_{\alpha d}^{B} = E_{\boldsymbol{y}_{dr}}\left[F_{\alpha d} | \boldsymbol{y}_{ds} \right], \quad F_{\alpha d} = h_\alpha(\boldsymbol{y}_d),$$

where $\boldsymbol{y}_{ds}$ and $\boldsymbol{y}_{dr}$ denote respectively sample and out-of-sample parts of $\boldsymbol{y}_d$.

EB method for poverty estimation

- **Empirical Best (EB) estimator:** Expectation calculated with respect to the distribution of $\boldsymbol{y}_{dr}|\boldsymbol{y}_{ds}$ with estimated unknown parameters.

- **Nested error linear model:**

$$Y_{dj} = \mathbf{x}_{dj}\boldsymbol{\beta} + u_d + e_{dj}, \quad j = 1, \ldots, N_d,\ d = 1, \ldots, D.$$
$$u_d \overset{iid}{\sim} N(0, \sigma_u^2), \quad e_{dj} \overset{iid}{\sim} N(0, \sigma_e^2).$$

- Distribution of $\boldsymbol{y}_d$:

$$\boldsymbol{y}_d \overset{ind}{\sim} N(\boldsymbol{\mu}_d, \mathrm{V}_d), \quad d = 1 \ldots, D,$$

where

$$\boldsymbol{\mu}_d = \boldsymbol{X}_d\boldsymbol{\beta} \text{ and } \mathrm{V}_d = \sigma_u^2 \mathbf{1}_{N_d}\mathbf{1}'_{N_d} + \sigma_e^2 I_{N_d}.$$

EB method for poverty estimation

- Decomposition in sample and out-of-sample:

$$\boldsymbol{\mu}_d = \begin{pmatrix} \boldsymbol{\mu}_{ds} \\ \boldsymbol{\mu}_{dr} \end{pmatrix}, \quad \mathbf{V}_d = \begin{pmatrix} \mathbf{V}_{ds} & \mathbf{V}_{dsr} \\ \mathbf{V}_{drs} & \mathbf{V}_{dr} \end{pmatrix}$$

- Distribution of $\boldsymbol{y}_{dr}$ given $\boldsymbol{y}_{ds}$:

$$\boldsymbol{y}_{dr} | \boldsymbol{y}_{ds} \sim N(\boldsymbol{\mu}_{dr|ds}, \mathbf{V}_{dr|ds}),$$

where

$$\boldsymbol{\mu}_{dr|ds} = \boldsymbol{\mu}_{dr} + \mathbf{V}_{drs}\mathbf{V}_{ds}^{-1}(\boldsymbol{y}_{ds} - \boldsymbol{\mu}_{ds}),$$
$$\mathbf{V}_{dr|ds} = \mathbf{V}_{dr} - \mathbf{V}_{drs}\mathbf{V}_{ds}^{-1}\mathbf{V}_{dsr}.$$

EB method for poverty estimation

- **Monte Carlo approximation of best estimator:**

(a) Generate L non-sample vectors $\boldsymbol{y}_{dr}^{(\ell)}$, $\ell = 1, \ldots, L$ from the conditional distribution of $\boldsymbol{y}_{dr}|\boldsymbol{y}_{ds}$.

(b) Attach the sample elements to form a population vector $\boldsymbol{y}_{d}^{(\ell)} = (\boldsymbol{y}_{ds}, \boldsymbol{y}_{dr}^{(\ell)})$, $\ell = 1, \ldots, L$.

(c) Calculate the poverty measure with each population vector $F_{\alpha d}^{(\ell)} = h_\alpha(\boldsymbol{y}_{d}^{(\ell)})$, $\ell = 1, \ldots, L$. Then take the average over the L Monte Carlo generations:

$$\hat{F}_{\alpha d}^{B} = E_{\boldsymbol{y}_{dr}}[F_{\alpha d}|\boldsymbol{y}_{ds}] \cong \frac{1}{L}\sum_{\ell=1}^{L} F_{\alpha d}^{(\ell)}.$$

Time mixed models

- Unit-level linear mixed models

$$y_{dtj} = \mathbf{x}_{dtj}\boldsymbol{\beta} + u_{1,d} + u_{2,dt} + w_{dtj}^{-1/2} e_{dtj}, \qquad \begin{array}{l} d = 1, \ldots, D, \\ t = 1, \ldots, m_d, \\ j = 1, \ldots, n_{dt}. \end{array} \tag{1}$$

where

(TM1) $u_{1,d}$ i.i.d. $N(0, \sigma_1^2)$, $(u_{2,d1}, \ldots, u_{2,dm_d})$ i.i.d. $AR(1; \sigma_2^2, \rho)$ and e_{dtj} i.i.d. $N(0, \sigma_0^2)$ are independent.

(TM2) $u_{1,d}$ i.i.d. $N(0, \sigma_1^2)$, $u_{2,dt}$ i.i.d. $N(0, \sigma_2^2)$ and e_{dtj} i.i.d. $N(0, \sigma_0^2)$ are independent.

Time mixed models

- Area-level linear mixed models

$$y_{dt} = \mathbf{x}_{dt}\boldsymbol{\beta} + u_{dt} + e_{dt}, \quad d = 1, \dots, D, \quad t = 1, \dots, m_d,$$

where

$\hookrightarrow$ y_{dt} is a direct estimator of the characteristic of interest and

$\hookrightarrow$ $\mathbf{x}_{dt}$ is a vector containing the population (aggregated) values of p auxiliary variables.

(TM3) $(u_{d1}, \dots, u_{2dm_d})$ i.i.d. $AR(1; \sigma_u^2, \rho)$ and $e_{dt} \overset{ind}{\sim} N(0, \sigma_{dt}^2)$ are independent.

(TM4) u_{dt} i.i.d. $N(0, \sigma_u^2)$ and $e_{dt} \overset{ind}{\sim} N(0, \sigma_{dt}^2)$ are independent.

Spatial mixed models

- Unit-level mixed models

$$y_{dj} = \mathbf{x}_{dj}\boldsymbol{\beta} + v_d + w_{dj}^{-1/2} e_{dj}, \quad d = 1, \ldots, D,\ j = 1, \ldots, n_d \tag{2}$$

where

(SM1) $(v_1, \ldots, v_D)$ i.i.d. $SAR(1; \sigma_v^2, \rho, \mathbf{P})$ and e_{dj} i.i.d. $N(0, \sigma_e^2)$ are independent.

- Area-level linear mixed models

$$y_d = \mathbf{x}_d\boldsymbol{\beta} + v_d + e_d, \quad d = 1, \ldots, D,$$

where
$\hookrightarrow$ y_d is a direct estimator of the characteristic of interest and
$\hookrightarrow$ $\mathbf{x}_d$ is a vector containing the population (aggregated) values of p auxiliary variables.

(SM2) $(v_1, \ldots, v_D)$ i.i.d. $SAR(1; \sigma_v^2, \rho, \mathbf{P})$ and $e_d \overset{ind}{\sim} N(0, \sigma_d^2)$
are independent.

spatio-temporal mixed models

- Area-level spatio-temporal linear mixed models

$$y_{dt} = \mathbf{x}_{dt}\boldsymbol{\beta} + u_{1d} + u_{2dt} + e_{dt}, \quad d = 1, \dots, D, \quad t = 1, \dots, T,$$

where

$\hookrightarrow$ y_{dt} is a direct estimator of the characteristic of interest and

$\hookrightarrow$ $\mathbf{x}_{dt}$ is a vector containing the population (aggregated) values of p auxiliary variables.

(STM1) $\{u_{1d}\}$, $\{u_{2dt}\}$ and $\{e_{dt}\}$ are independent with distributions $\{u_{1d}\}_{d=1}^{D} \sim SAR(1)$, $\{u_{2dt}\}$ i.i.d $N(0, \sigma_2^2)$ and $e_{dt} \sim N(0, \sigma_{dt}^2)$.

(STM2) $\{u_{1d}\}$, $\{u_{2dt}\}$ y $\{e_{dt}\}$ are independent with distributions $\{u_{1d}\}_{d=1}^{D} \sim SAR(1)$, $\{u_{2dt}\}_{t=1}^{T}$ i.i.d $AR(1)$ and $e_{dt} \sim N(0, \sigma_{dt}^2)$.

Semiparametric Fay-Herriot model

Extension of the Fay-Herriot model allowing non linearity in the relationship between $\boldsymbol{\theta}$ and $\mathbf{X}$

- $\mathbf{x}_1$: let us consider one covariate for simplicity
- $\tilde{m}(\mathbf{x}_1)$: nonparametric model with one covariate (x_1), where $\tilde{m}(\cdot)$ is unknown bu sufficiently well approximated by the function $m(\mathbf{x}_1; \boldsymbol{\eta}; \boldsymbol{\gamma})$
- $m(\mathbf{x}_1; \boldsymbol{\eta}; \boldsymbol{\gamma}) = \eta_0 + \eta_1 \mathbf{x}_1 + \ldots + \eta_p \mathbf{x}_1^p + \sum_{k=1}^{K} \gamma_k (\mathbf{x}_1 - \kappa_k)_+^p$
- $\boldsymbol{\eta}$: is the $(p+1) \times 1$ vector of the coefficient of the polynomial function
- $\boldsymbol{\gamma}$: is the $k \times 1$ vector of the truncated polynomial spline basis (P-spline with degree p)
- $\kappa_k, k = 1, \ldots, K$: is a set of fixed knots
- define $\mathbf{X}_1$ and $\mathbf{Z}$ as the matrix of polynomial values and truncated polynomial spline respectively:

$$\mathbf{X}_1 = \begin{bmatrix} 1 & x_{11} & \cdots & x_{11}^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1m} & \cdots & x_{1m}^p \end{bmatrix}, \mathbf{Z} = \begin{bmatrix} (x_{11} - \kappa_1)_+^p & \cdots & (x_{11} - \kappa_K)_+^p \\ \vdots & \ddots & \vdots \\ (x_{1m} - \kappa_1)_+^p & \cdots & (x_{1m} - \kappa_K)_+^p \end{bmatrix}$$

Semiparametric Fay-Herriot model

Mixed model representation of the semiparametric Fay-Herriot model

$$\boldsymbol{\theta} = \begin{bmatrix} \mathbf{X} \\ \mathbf{X}_1 \end{bmatrix} [\boldsymbol{\beta}, \boldsymbol{\eta}] + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{D}\mathbf{u} + \boldsymbol{\epsilon}$$

- $\boldsymbol{\alpha} \rightarrow [\boldsymbol{\beta}, \boldsymbol{\eta}]$
- $\mathbf{A} \rightarrow [\mathbf{X}, \mathbf{X}_1]$

i.e. $\mathbf{A} = [\mathbf{x}_1, \ldots, \mathbf{x}_q]$, $\mathbf{X}_1$ as expressed above, then

$$\mathbf{A} = \begin{bmatrix} 1 & x_{21} & \ldots & x_{q1} & x_{11} & \ldots & x_{11}^p \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{2m} & \ldots & x_{qm} & x_{1m} & \ldots & x_{1m}^p \end{bmatrix}, \boldsymbol{\alpha} = [\eta_0, \beta_2, \ldots, \beta_q, \eta_1, \ldots, \eta_p]^T$$

Semiparametric Fay-Herriot model

$$\boldsymbol{\theta} = \mathbf{A}\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{D}\mathbf{u} + \boldsymbol{\epsilon}$$

Semiparametric Fay-Herriot model

Model-based estimator of the small area mean (or total):

$$\hat{\boldsymbol{\theta}}^E(\hat{\boldsymbol{\psi}}) = \mathbf{A}\hat{\boldsymbol{\alpha}}(\hat{\boldsymbol{\psi}}) + \hat{\boldsymbol{\Lambda}}(\hat{\boldsymbol{\psi}})[\hat{\boldsymbol{\theta}} - \mathbf{A}\hat{\boldsymbol{\alpha}}(\hat{\boldsymbol{\psi}})]$$

- $\boldsymbol{\Lambda}(\psi) = (\mathbf{Z}\boldsymbol{\Sigma}_\gamma\mathbf{Z}^T + \mathbf{D}\boldsymbol{\Sigma}_u\mathbf{D}^T)\boldsymbol{\Sigma}^{-1}(\psi)$
- $\boldsymbol{\alpha}(\boldsymbol{\psi}) = (\mathbf{A}^T\boldsymbol{\Sigma}^{-1}(\psi)\mathbf{A})^{-1}\mathbf{A}^T\boldsymbol{\Sigma}^{-1}(\psi)\hat{\boldsymbol{\theta}}$
- $\boldsymbol{\Sigma}(\boldsymbol{\psi}) = \mathbf{Z}\boldsymbol{\Sigma}_\gamma\mathbf{Z}^T + \mathbf{D}\boldsymbol{\Sigma}_u\mathbf{D}^T + \mathbf{R}$
- $\boldsymbol{\Sigma}_u = \sigma_u^2\mathbf{I}_m$
- $\boldsymbol{\Sigma}_\gamma = \sigma_\gamma^2\mathbf{I}_K$
- $\boldsymbol{\psi} = [\sigma_u^2, \sigma_\gamma^2]^T$

REMARK: $\hat{\cdot}$ means that the unknown parameter $\cdot$ has been replaced by REML estimator
REMARK: mean square error estimator for $\hat{\boldsymbol{\theta}}^E$ can be obtained by Taylor approximation (Prasad and Rao, 1990 and Opsomer et al., 2006) or by a bootstrap approach

M-quantile models

- With regression models we model the mean of the variable of interest (y) given the covariates (x)

- A more complete picture is offered by modeling not only the mean of y given x, but also the quantiles.

- Examples include the median, the 25th, 75th percentiles.

- This is known as quantile regression

- An M-quantile regression model for quantile q is

$$Q_q(\boldsymbol{y}|\boldsymbol{X}) = \boldsymbol{X}\boldsymbol{\beta}(q)$$

where q is a-priori chosen.

M-quantile models

- Estimate of $\boldsymbol{\beta}(q)$ is obtained via Iterative Weighted Least Squares:

$$\widehat{\boldsymbol{\beta}}(q) = (\boldsymbol{X}^t \mathbf{W} \boldsymbol{X})^{-1} \boldsymbol{X}^t \mathbf{W} \boldsymbol{y}$$

- $\mathbf{W}$ is an $n \times n$ diagonal weighting matrix that depends on both the influence function and the quantile we are modeling

- Central Idea: Area effects can be described by estimating an area specific q value ($\widehat{\theta}_d$) for each area (group) of a hierarchical dataset

- Estimate the area specific target parameter by fitting an M-quantile model for each area at $\widehat{\theta}_d$

- A mixed model uses random effects u_d to capture the dissimilarity between groups. M-quantile models attempt to capture this dissimilarity via the group-specific M-quantile coefficients $\widehat{\theta}_d$

Estimation of cumulative distribution functions

- Estimation of the distribution function of income will be performed using both M-quantile and random effects models (see Chambers and Dunstan 1986 ; Rao, Kovar and Mantel 1990).

- The CDF estimator can be further used for estimating other quantiles of the small area distribution function of the variable of interest.

- This is achieved by integrating the CDF estimator

$$\int_{-\infty}^{q} t \, d\widehat{F}_d(t)$$

Small Area Estimation by Borrowing Strength over Space

- In applications involving economic, environmental and epidemiological data observations that are spatially close may be more alike than observations that are further apart .

- This creates a type of spatial dependency or spatial association in the data that invalidates the assumption of independent and identically distributed (iid) observations used by conventional regression models.

- One approach to accounting for spatial correlation in the data is offered by specifying models with spatially correlated errors (Anselin 1992; Cressie 1993).

- Small area literature suggests that prediction of small area parameters may be improved by borrowing strength over space (Saei & Chambers 2003; Singh *et al.* 2005; Petrucci & Salvati 2006; Pratesi & Salvati 2008; 2009; Molina *et al.* 2008).

Global Vs. Local Models for Modeling Spatial Dependency

- Regression models with spatially correlated errors are global models i.e. they assume that the relationship we are modelling holds everywhere in the study area.

- Another approach to modelling a spatially non-stationary process is offered via Geographically Weighted Regression (GWR) (Brunsdon *et al.* 1996; Fotheringham *et al.* 1997).

- GWR models attempt to capture the spatial association in the data by allowing local, rather than global parameters, to be estimated.

GWR Models

- Assume that we have n observations on $(y_j, \mathbf{x}_j)$ at a set of Locations (u_j).

- A GWR model is defined as follows

$$y_j = \mathbf{x}_j^T \boldsymbol{\beta}(u_j) + e(u_j)$$

- GWR models allow for local rather than global parameters to be estimated and will produce estimated local surfaces of the relationship between y and x.

- GWR models work by assuming that observed data near to location j will have a greater influence on the estimation of $\boldsymbol{\beta}(u_j)$ than observations farther from j.

- Weighted Least Squares (WLS) is used for estimating the GWR parameters.

Extend GWR to outlier robust GWR

- M-quantile Geographically Weighted Models (Salvati, Tzavidis, Pratesi & Chambers 2007.

- We first propose a robust GWR model namely an M-quantile GWR model. This is a locally robust to outliers model

- With this model we attempt to model locally the different quantiles of the conditional distribution accounting at the same time for the spatial non-stationarity in the data

- For estimating the parameters of the M-quantile GWR model we use an Iterative Weighted Least Squares algorithm

M-quantile GWR

- The M-quantile GWR model is defined as follows

$$Q_q(y|\mathbf{X}, u) = \mathbf{X}^T \beta(u; q)$$

- The model parameters $\beta(u_i; q)$ are estimated by solving

$$\sum_{l=1}^{L} w(u_l, u) \sum_{i=1}^{n_l} \psi_q \left\{ y_{il} - \mathbf{x}_{il}^T \beta(u_i; q) \right\} \mathbf{x}_{il} = 0$$

- Solution depends on distance function, bandwidth and influence function
- Estimates of $\beta(u_i; q)$'s via IWLS

$$\hat{\beta}(u; q) = (\mathbf{X}^T \mathbf{W}^* \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^* \mathbf{Y}$$

M-quantile GWR SAE

- Achieved via an extension of the M-quantile SAE algorithm
- Step 1: Estimate unit level M-quantile coefficients $\hat{\theta}_{ij}$, using M-quantile GWR. $\hat{\theta}_{ij}$'s are accounting for the spatial structure
- Step 2: Recognize the area structure and estimate area M-quantile coefficients, $\hat{\theta}_j$, using $\hat{\theta}_{ij}$
- Step 3: Estimate the area target parameter using an M-quantile GWR model for each area at $\hat{\theta}_j$

$$Q_{\hat{\theta}_j}(y|\mathbf{X}, u) = \mathbf{X}^T \hat{\beta}(u; \hat{\theta}_j)$$

M-quantile GWR SAE

- The 'naïve' small area estimator is

$$\hat{m}_j^{MQGWR} = N_j^{-1}\{\sum_{i\in s_j} y_{ij} + \sum_{i\in r_j} \mathbf{x}_{ij}^T \hat{\beta}(u_i;\hat{\theta}_j)\}$$

- A bias-corrected small area estimator is

$$\hat{m}_j^{MQGWR/CD} = N_j^{-1}\{\sum_{i\in s_j} y_{ij} + \sum_{i\in r_j} \mathbf{x}_{ij}^T \hat{\beta}(u_i;\hat{\theta}_j) + \frac{N_j - n_j}{n_j}\sum_{i\in s_j}[y_{ij} - \hat{y}_{ij}]\}$$

Reflections on the use of M-quantile SAE models

- Less parametric - No assumptions on the random effects
- Outlier robust
- Approximates $Q_q(y|\mathbf{X})$ with a local linear function
- Larger number of parameters to be estimated compared to SAR
- Improved estimation for in and out of sample area estimation

Motivating the use of nonparametric models in SAE

- Random effects models are higly structured and parametric
- M-quantile models impose fewer assumptions about the structure of the data and are less parametric allowing for outlier robust inference
- All models, however, assume linearity between y and x
- What if the linearity assumption fails?
- Two possibilities:
  - Nonparametric random effects models(Opsomer et al., 2008)
  - Nonparametric M-quantile models (Pratesi et al., 2008)

Nonparametric M-quantile SAE models

- Proposed by Pratesi et al.(2008) for relaxing the linearity of the M-quantile SAE model

$$Q_q(y|\mathbf{X}) = m(\mathbf{X}; \boldsymbol{\beta}(q), \boldsymbol{\nu}(q))$$

  where $m(\mathbf{X}; \boldsymbol{\beta}(q), \boldsymbol{\nu}(q)) = \mathbf{X}\boldsymbol{\beta}(q) + \boldsymbol{\delta}^T \boldsymbol{\nu}(q)$ where $\boldsymbol{\nu}(q)$ denotes the spline function at quantile q
- Borrow strength over space in the spirit of Opsomer et al.(2008): use a bivariate spline on the spatial coordinates
- Naive and biased adjusted small area estimators under the nonparametric M-quantile model can be defined

$$\hat{m}_j^{NPMQ} = N_j^{-1}\Big(\sum_{i \in s_j} y_{ij} + \sum_{i \in r_j} \hat{y}_{ij}\Big)$$

References

Chambers, R. and Dunstan, R. (1986). Estimating distribution function from survey data, Biometrika, **73**, 597-604.

Chambers, R. and Tzavidis, N. (2006). M-quantile Models for Small Area Estimation. *Biometrika*, **93**, 255–268.

Foster, J., Greer, J. and Thorbecke, E. (1984). A class of decomposable poverty measures, *Econometrica*, **52**, 761–766.

Pfeffermann, D., Terryn, B. and Moura, F. (2007). *Bayesian small area estimation of literacy under a two part random effects model*, Presentation at SAE2007 Conference held in Pisa.

Rao, J.N.K., Kovar, J.G. and Mantel, H.J. (1990). On estimating distribution functions and quantiles from survey data using auxiliary information. Biometrika, **2**, 365-375.

Salvati, N., Tzavidis, N., Pratesi, M. and Chambers, R. (2007). Small Area Estimation Via M-quantile Geographically Weighted Regression. [*Paper submitted for publication. A copy is available upon request]*.

Tzavidis, N., Salvati, N., Pratesi, M. and Chambers, R. (2007). M-quantile models for poverty mapping. *Statistical Methods & Applications*, 17, 393-411.