An application of gamma mixed models to small area estimation

Tomáš Hobza

Department of Mathematics

Czech Technical University in Prague, Czech Republic

Based on joined work with

Domingo Morales and **Yolanda Marhuenda**

University of Miguel Hernández, Elche, Spain

SPMS2018, 22.6.2018, Dobřichovice

Outline

Introduction

- U **finite** population of size N.
- D - domains
- N_d - population size, $\quad d = 1, \ldots, D$
- Variable of interest Y.
- y_{dj} value of Y in unit j from domain d

- **Target:** to estimate additive parameters of Y in the D domains/areas.

Introduction

- U **finite** population of size N.
- D - domains
- N_d - population size, $\quad d = 1, \ldots, D$
- Variable of interest Y.
- y_{dj} value of Y in unit j from domain d
- **Target:** to estimate additive parameters of Y in the D domains/areas.

Introduction

- Our parameter of interest is

$$\delta_d = \frac{1}{N_d} \sum_{j=1}^{N_d} h(y_{dj}),$$

  where h is a known measurable function.
- For $h(y) = y$ we obtain the area mean income

$$\overline{Y}_d = \frac{1}{N_d} \sum_{j=1}^{N_d} y_{dj}.$$

- For $h(y) = I(y < z)$ we obtain the area poverty proportions

$$p_d = \frac{1}{N_d} \sum_{j=1}^{N_d} I\left(y_{dj} < z\right).$$

Introduction

- We have a sample $S \subset U$ of size n drawn from the whole population.
- $S_d = S \cap U_d$ sub-sample from domain d of size n_d.

Direct estimates of $\delta_d = \frac{1}{N_d} \sum_{j=1}^{N_d} h(y_{dj})$ are

$$\widehat{\delta}_d^{dir} = \frac{1}{\widehat{N}_d} \sum_{j \in S_d} w_{dj} h(y_{dj}), \qquad \widehat{N}_d = \sum_{j \in S_d} w_{dj}$$

where w_{dj} are the calibrated sampling weights.

Introduction

- We have a sample $S \subset U$ of size n drawn from the whole population.
- $S_d = S \cap U_d$ sub-sample from domain d of size n_d.

Direct estimates of $\delta_d = \frac{1}{N_d} \sum_{j=1}^{N_d} h(y_{dj})$ are

$$\widehat{\delta}_d^{dir} = \frac{1}{\widehat{N}_d} \sum_{j \in S_d} w_{dj} h(y_{dj}), \qquad \widehat{N}_d = \sum_{j \in S_d} w_{dj}$$

where w_{dj} are the calibrated sampling weights.

Introduction

- Under SRS without replacement within each area,

$$w_{dj} = \frac{N_d}{n_d},\ \forall j \in S_d \quad \Rightarrow \quad \widehat{\delta}_d^{dir} = \frac{1}{n_d} \sum_{j \in S_d} h(y_{dj}).$$

- **Problem:** n_d **small** for some d.
- **Small area/domain:** subset of the population that is target of inference and for which the direct estimator does not have enough precision.
- What does "enough precision" means? Some National Statistical Offices (Spain) allow a maximum CV of 20%.

Introduction

- Under SRS without replacement within each area,

$$w_{dj} = \frac{N_d}{n_d},\ \forall j \in S_d \quad \Rightarrow \quad \widehat{\delta}_d^{dir} = \frac{1}{n_d} \sum_{j \in S_d} h(y_{dj}).$$

- **Problem:** n_d **small** for some d.
- **Small area/domain:** subset of the population that is target of inference and for which the direct estimator does not have enough precision.
- What does "enough precision" means? Some National Statistical Offices (Spain) allow a maximum CV of 20%.

Introduction

- Under SRS without replacement within each area,

$$w_{dj} = \frac{N_d}{n_d},\ \forall j \in S_d \quad \Rightarrow \quad \widehat{\delta}_d^{dir} = \frac{1}{n_d} \sum_{j \in S_d} h(y_{dj}).$$

- **Problem:** n_d **small** for some d.
- **Small area/domain:** subset of the population that is target of inference and for which the direct estimator does not have enough precision.
- What does "enough precision" means? Some National Statistical Offices (Spain) allow a maximum CV of 20%.

Unit level gamma mixed model

- The distribution of the target variable y_{dj}, conditioned to the random effect v_d is for $j = 1, \ldots, N_d$

$$y_{dj}|_{v_d} \sim \text{Gamma}\big(\nu_{dj},\, \alpha_{dj} = \frac{\nu_{dj}}{\mu_{dj}}\big), \quad \nu_{dj} = a_{dj}\varphi.$$

- For the inverse of the mean parameter, we assume

$$g(\mu_{dj}) = \frac{1}{\mu_{dj}} = \mathbf{x}_{dj}^T\boldsymbol{\beta} + \phi v_d,$$

where

  - $\{v_d : \ d = 1, \ldots, D\}$ are i.i.d. $N(0, 1)$
  - y_{dj}'s are independent conditioned to $\mathbf{v}$.

- The vector of unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}, \phi, \varphi)$ is estimated by maximizing the Laplace approximation of the log-likelihood.

Unit level gamma mixed model

- The distribution of the target variable y_{dj}, conditioned to the random effect v_d is for $j = 1, \ldots, N_d$

$$y_{dj}|_{v_d} \sim \text{Gamma}\big(\nu_{dj},\, \alpha_{dj} = \frac{\nu_{dj}}{\mu_{dj}}\big), \quad \nu_{dj} = a_{dj}\varphi.$$

- For the inverse of the mean parameter, we assume

$$g(\mu_{dj}) = \frac{1}{\mu_{dj}} = \mathbf{x}_{dj}^T\boldsymbol{\beta} + \phi v_d,$$

where

  - $\{v_d : \ d = 1, \ldots, D\}$ are i.i.d. $N(0, 1)$
  - y_{dj}'s are independent conditioned to $\mathbf{v}$.

- The vector of unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}, \phi, \varphi)$ is estimated by maximizing the Laplace approximation of the log-likelihood.

Unit level gamma mixed model

- The distribution of the target variable y_{dj}, conditioned to the random effect v_d is for $j = 1, \ldots, N_d$

$$y_{dj}|_{v_d} \sim \text{Gamma}\big(\nu_{dj},\, \alpha_{dj} = \frac{\nu_{dj}}{\mu_{dj}}\big), \quad \nu_{dj} = a_{dj}\varphi.$$

- For the inverse of the mean parameter, we assume

$$g(\mu_{dj}) = \frac{1}{\mu_{dj}} = \mathbf{x}_{dj}^T\boldsymbol{\beta} + \phi v_d,$$

where

  - $\{v_d : \ d = 1, \ldots, D\}$ are i.i.d. $N(0, 1)$
  - y_{dj}'s are independent conditioned to $\mathbf{v}$.

- The vector of unknown parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}, \phi, \varphi)$ is estimated by maximizing the Laplace approximation of the log-likelihood.

Empirical best predictor

- Our parameter of interest is

$$\delta_d = \frac{1}{N_d} \sum_{j=1}^{N_d} h(y_{dj}).$$

- Let us denote by S_d and R_d the sets of sampled and non-sampled individuals in domain d
- Best predictor (BP) of δ_d is

$$\hat{\delta}_d = \hat{\delta}_d(\theta) = \frac{1}{N_d} \Big[\sum_{j \in S_d} h(y_{dj}) + \sum_{j \in R_d} E_{\theta}[h(y_{dj}) | \mathbf{y}_s] \Big].$$

- We would need a census file with all the **x** variables
- Might be overcome if all the **x** variables are categorical

Empirical best predictor

- Our parameter of interest is

$$\delta_d = \frac{1}{N_d} \sum_{j=1}^{N_d} h(y_{dj}).$$

- Let us denote by S_d and R_d the sets of sampled and non-sampled individuals in domain d
- Best predictor (BP) of δ_d is

$$\hat{\delta}_d = \hat{\delta}_d(\boldsymbol{\theta}) = \frac{1}{N_d} \Big[\sum_{j \in S_d} h(y_{dj}) + \sum_{j \in R_d} E_{\boldsymbol{\theta}}[h(y_{dj}) | \mathbf{y}_s] \Big].$$

- We would need a census file with all the **x** variables
- Might be overcome if all the **x** variables are categorical

Empirical best predictor

- Our parameter of interest is

$$\delta_d = \frac{1}{N_d} \sum_{j=1}^{N_d} h(y_{dj}).$$

- Let us denote by S_d and R_d the sets of sampled and non-sampled individuals in domain d
- Best predictor (BP) of δ_d is

$$\hat{\delta}_d = \hat{\delta}_d(\boldsymbol{\theta}) = \frac{1}{N_d} \Big[\sum_{j \in S_d} h(y_{dj}) + \sum_{j \in R_d} E_{\boldsymbol{\theta}}[h(y_{dj}) | \mathbf{y}_s] \Big].$$

- We would need a census file with all the **x** variables
- Might be overcome if all the **x** variables are categorical

Empirical best predictor

- Suppose that the covariates are categorical such that

$$\mathbf{x}_{dj} \in \{\mathbf{z}_1, \ldots, \mathbf{z}_K\}.$$

- Then

$$\sum_{j \in R_d} E_{\boldsymbol{\theta}}[h(y_{dj})|\mathbf{y}_s] = \sum_{k=1}^{K} w_{dk} E_{\boldsymbol{\theta}}[h(y_{dk})|\mathbf{y}_s],$$

where $y_{dk} \sim Gamma\left(\nu_{dk}, \frac{\nu_{dk}}{\mu_{dk}}\right)$,

$$\mu_{dk} = \mu_{dk}(\boldsymbol{\theta}) = \left(\mathbf{z}_k^T \boldsymbol{\beta} + \phi v_d\right)^{-1}$$

and

$$w_{dk} = \#\{j \in R_d : \mathbf{x}_{dj} = \mathbf{z}_k\}$$

is the size of the covariate class $\mathbf{z}_k$ at R_d (available from external data sources).

Empirical best predictor

- Suppose that the covariates are categorical such that

$$\mathbf{x}_{dj} \in \{\mathbf{z}_1, \ldots, \mathbf{z}_K\}.$$

- Then

$$\sum_{j \in R_d} E_{\boldsymbol{\theta}}[h(y_{dj})|\mathbf{y}_s] = \sum_{k=1}^{K} w_{dk} E_{\boldsymbol{\theta}}[h(y_{dk})|\mathbf{y}_s],$$

where $y_{dk} \sim Gamma\left(\nu_{dk}, \frac{\nu_{dk}}{\mu_{dk}}\right)$,

$$\mu_{dk} = \mu_{dk}(\boldsymbol{\theta}) = \left(\mathbf{z}_k^T \boldsymbol{\beta} + \phi v_d\right)^{-1}$$

and

$$w_{dk} = \#\{j \in R_d : \mathbf{x}_{dj} = \mathbf{z}_k\}$$

is the size of the covariate class $\mathbf{z}_k$ at R_d (available from external data sources).

Empirical best predictor

- In this categorical setup the BP of δ_d is

$$\hat{\delta}_d^{BP}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}[\delta_d|\mathbf{y}_s] = \frac{1}{N_d}\Big[\sum_{j\in S_d} h(y_{dj}) + \sum_{k=1}^{K} w_{dk} E_{\boldsymbol{\theta}}[h(y_{dk})|\mathbf{y}_s]\Big],$$

  where

$$E_{\boldsymbol{\theta}}[h(y_{dk})|\mathbf{y}_s]$$

  must be approximated numerically.

- The EBP of δ_d is then obtained as

$$\hat{\delta}_d^{EBP} = \hat{\delta}_d^{BP}(\hat{\boldsymbol{\theta}})\,.$$

Empirical best predictor

- In this categorical setup the BP of δ_d is

$$\hat{\delta}_d^{BP}(\boldsymbol{\theta}) = E_{\boldsymbol{\theta}}[\delta_d|\mathbf{y}_s] = \frac{1}{N_d}\Big[\sum_{j\in S_d} h(y_{dj}) + \sum_{k=1}^{K} w_{dk} E_{\boldsymbol{\theta}}[h(y_{dk})|\mathbf{y}_s]\Big],$$

  where

$$E_{\boldsymbol{\theta}}[h(y_{dk})|\mathbf{y}_s]$$

  must be approximated numerically.

- The EBP of δ_d is then obtained as

$$\hat{\delta}_d^{EBP} = \hat{\delta}_d^{BP}(\hat{\boldsymbol{\theta}})\,.$$

PLUG-IN estimator

The plug-in estimator of δ_d is

$$\tilde{\delta}_d = \tilde{\delta}_d(\hat{\boldsymbol{\theta}}) = \frac{1}{N_d}\Big[\sum_{j \in S_d} h(y_{dj}) + \sum_{k=1}^{K} w_{dk}\, h(\tilde{\mu}_{dk})\Big],$$

where

$$\tilde{\mu}_{dk} = \left(\mathbf{z}_k^T \hat{\boldsymbol{\beta}} + \hat{\phi}\hat{v}_d\right)^{-1}.$$

- PROBLEM: for the function $h(y) = I(y < z)$

$$h(\tilde{\mu}_{dk}) = I(\tilde{\mu}_{dk} < z) = 0$$

PLUG-IN estimator

The plug-in estimator of δ_d is

$$\tilde{\delta}_d = \tilde{\delta}_d(\hat{\boldsymbol{\theta}}) = \frac{1}{N_d}\Big[\sum_{j \in S_d} h(y_{dj}) + \sum_{k=1}^{K} w_{dk}\, h(\tilde{\mu}_{dk})\Big],$$

where

$$\tilde{\mu}_{dk} = \left(\mathbf{z}_k^T \hat{\boldsymbol{\beta}} + \hat{\phi}\hat{v}_d\right)^{-1}.$$

- PROBLEM: for the function $h(y) = I(y < z)$

$$h(\tilde{\mu}_{dk}) = I(\tilde{\mu}_{dk} < z) = 0$$

Marginal predictor

Let us consider the predicted marginal distribution of y_{dk}, i.e. the p.d.f. and d.f. of

$$Gamma\left(\widehat{\nu}_{dk}, \frac{\widehat{\nu}_{dk}}{\tilde{\mu}_{dk}}\right).$$

The marginal predictor of δ_d is

$$\hat{\delta}_d^{MAR} = \frac{1}{N_d}\Big[\sum_{j\in S_d} h(y_{dj}) + \sum_{k=1}^{K} w_{dk} E[h(y_{dk})|\hat{\nu}_{dk}, \tilde{\mu}_{dk}]\Big].$$

- For $h(y) = y$ we get

$$E[h(y_{dk})|\hat{\nu}_{dk}, \tilde{\mu}_{dk}] = \int_0^\infty y f(y|\hat{\nu}_{dk}, \tilde{\mu}_{dk})\, dy = \tilde{\mu}_{dk}.$$

- For the function $h(y) = I(y < z)$

$$E[h(y_{dk})|\hat{\nu}_{dk}, \tilde{\mu}_{dk}] = \int_0^z f(y|\hat{\nu}_{dk}, \tilde{\mu}_{dk})\, dy = F_{\hat{\nu}_{dk}, \tilde{\mu}_{dk}}(z).$$

Marginal predictor

Let us consider the predicted marginal distribution of y_{dk}, i.e. the p.d.f. and d.f. of

$$Gamma\left(\widehat{\nu}_{dk}, \frac{\widehat{\nu}_{dk}}{\tilde{\mu}_{dk}}\right).$$

The marginal predictor of δ_d is

$$\hat{\delta}_d^{MAR} = \frac{1}{N_d}\Big[\sum_{j\in S_d} h(y_{dj}) + \sum_{k=1}^{K} w_{dk} E[h(y_{dk})|\hat{\nu}_{dk}, \tilde{\mu}_{dk}]\Big].$$

- For $h(y) = y$ we get

$$E[h(y_{dk})|\hat{\nu}_{dk}, \tilde{\mu}_{dk}] = \int_0^{\infty} y f(y|\hat{\nu}_{dk}, \tilde{\mu}_{dk})\, dy = \tilde{\mu}_{dk}.$$

- For the function $h(y) = I(y < z)$

$$E[h(y_{dk})|\hat{\nu}_{dk}, \tilde{\mu}_{dk}] = \int_0^{z} f(y|\hat{\nu}_{dk}, \tilde{\mu}_{dk})\, dy = F_{\hat{\nu}_{dk}, \tilde{\mu}_{dk}}(z).$$

Marginal predictor

Let us consider the predicted marginal distribution of y_{dk}, i.e. the p.d.f. and d.f. of

$$Gamma\left(\widehat{\nu}_{dk}, \frac{\widehat{\nu}_{dk}}{\tilde{\mu}_{dk}}\right).$$

The marginal predictor of δ_d is

$$\hat{\delta}_d^{MAR} = \frac{1}{N_d}\Big[\sum_{j\in S_d} h(y_{dj}) + \sum_{k=1}^{K} w_{dk} E[h(y_{dk})|\hat{\nu}_{dk}, \tilde{\mu}_{dk}]\Big].$$

- For $h(y) = y$ we get

$$E[h(y_{dk})|\hat{\nu}_{dk}, \tilde{\mu}_{dk}] = \int_0^\infty y f(y|\hat{\nu}_{dk}, \tilde{\mu}_{dk})\, dy = \tilde{\mu}_{dk}.$$

- For the function $h(y) = I(y < z)$

$$E[h(y_{dk})|\hat{\nu}_{dk}, \tilde{\mu}_{dk}] = \int_0^z f(y|\hat{\nu}_{dk}, \tilde{\mu}_{dk})\, dy = F_{\hat{\nu}_{dk}, \tilde{\mu}_{dk}}(z).$$

Mean square error of predictors

Bootstrap estimator of MSE:

1) Fit the model to the sample and calculate $\hat{\boldsymbol{\theta}}$.

2) Repeat B times ($b = 1, \ldots, B$):
   a) Generate bootstrap population from the assumed model with the estimated $\hat{\boldsymbol{\theta}}$
   b) Calculate the true quantity $\delta_d^{*(b)}$
   c) Extract bootstrap sample, calculate $\hat{\boldsymbol{\theta}}^{*(b)}$ and the predictor $\hat{\delta}_d^{*(b)}$.

3) Output:

$$mse^*(\hat{\mu}_d) = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{\delta}_d^{*(b)} - \delta_d^{*(b)}\right)^2.$$

Simulation experiment

Target: to investigate the behaviour of the EBP and Marginal predictor.

Population generation

- Take $D = 30$, $N_d = 1\,000$ and $n_d \in \{25, 50, 75, 100\}$,
- For $d = 1, \ldots, D$ and $j = 1, \ldots, N_d$ generate regressors

$$(x_{dj1}, x_{dj2}) \in \{(0,0), (0,1), (1,0)\}$$

with probabilities equal to 0.3, 0.2 and 0.5, respectively.

It may represent belonging of the concrete individuum to one of three possible classes (e.g. inactive, unemployed and employed)

Simulation experiment

Target: to investigate the behaviour of the EBP and Marginal predictor.

Population generation

- Take $D = 30$, $N_d = 1\,000$ and $n_d \in \{25, 50, 75, 100\}$,
- For $d = 1, \ldots, D$ and $j = 1, \ldots, N_d$ generate regressors

$$(x_{dj1}, x_{dj2}) \in \{(0,0), (0,1), (1,0)\}$$

with probabilities equal to 0.3, 0.2 and 0.5, respectively.

It may represent belonging of the concrete individuum to one of three possible classes (e.g. inactive, unemployed and employed)

Simulation experiment

- Take $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2) = (0.8, -0.15, 0.2)$, $\phi = 0.1$ and $\varphi = 2.5$
- Generate $v_d \sim N(0, 1)$, $d = 1, \ldots, D$
- Generate the target variable as follows:

$$y_{dj} \sim \text{Gamma}\left(\nu_{dj}, \frac{\nu_{dj}}{\mu_{dj}}\right),$$

where

$$\mu_{dj} = \left(\beta_0 + x_{dj1}\beta_1 + x_{dj2}\beta_2 + \phi v_d\right)^{-1}, \quad \nu_{dj} = a_{dj}\varphi.$$

Simulation experiment

Steps of the simulation are:

1. **Repeat** $K = 1000$ **times** $(k = 1, \ldots, K)$
   1.1. Generate the population as described.
   1.2. Calculate the true values

$$p_d^{(k)} = \frac{1}{N_d} \sum_{j=1}^{N_d} I\left(y_{dj}^{(k)} < z\right)$$

   1.3. Select a simple random sample S_d (without replacement) of size n_d.
   1.4. Calculate:
   ✓ EBP ✓ MAR

2. **Output:** for each $\widehat{p}_d \in \{EBP, MAR\}$

$$B_d = \frac{1}{K} \sum_{k=1}^{K} (\widehat{p}_d^{(k)} - p_d^{(k)}), \quad E_d = \frac{1}{K} \sum_{k=1}^{K} (\widehat{p}_d^{(k)} - p_d^{(k)})^2.$$

Simulation experiment

Steps of the simulation are:

1. **Repeat** $K = 1000$ **times** $(k = 1, \ldots, K)$
   1.1. Generate the population as described.
   1.2. Calculate the true values

$$p_d^{(k)} = \frac{1}{N_d} \sum_{j=1}^{N_d} I\left(y_{dj}^{(k)} < z\right)$$

   1.3. Select a simple random sample S_d (without replacement) of size n_d.
   1.4. Calculate:
   ✓ EBP ✓ MAR

2. **Output:** for each $\widehat{p}_d \in \{EBP, MAR\}$

$$B_d = \frac{1}{K} \sum_{k=1}^{K} (\widehat{p}_d^{(k)} - p_d^{(k)}), \quad E_d = \frac{1}{K} \sum_{k=1}^{K} (\widehat{p}_d^{(k)} - p_d^{(k)})^2.$$

Simulation experiment

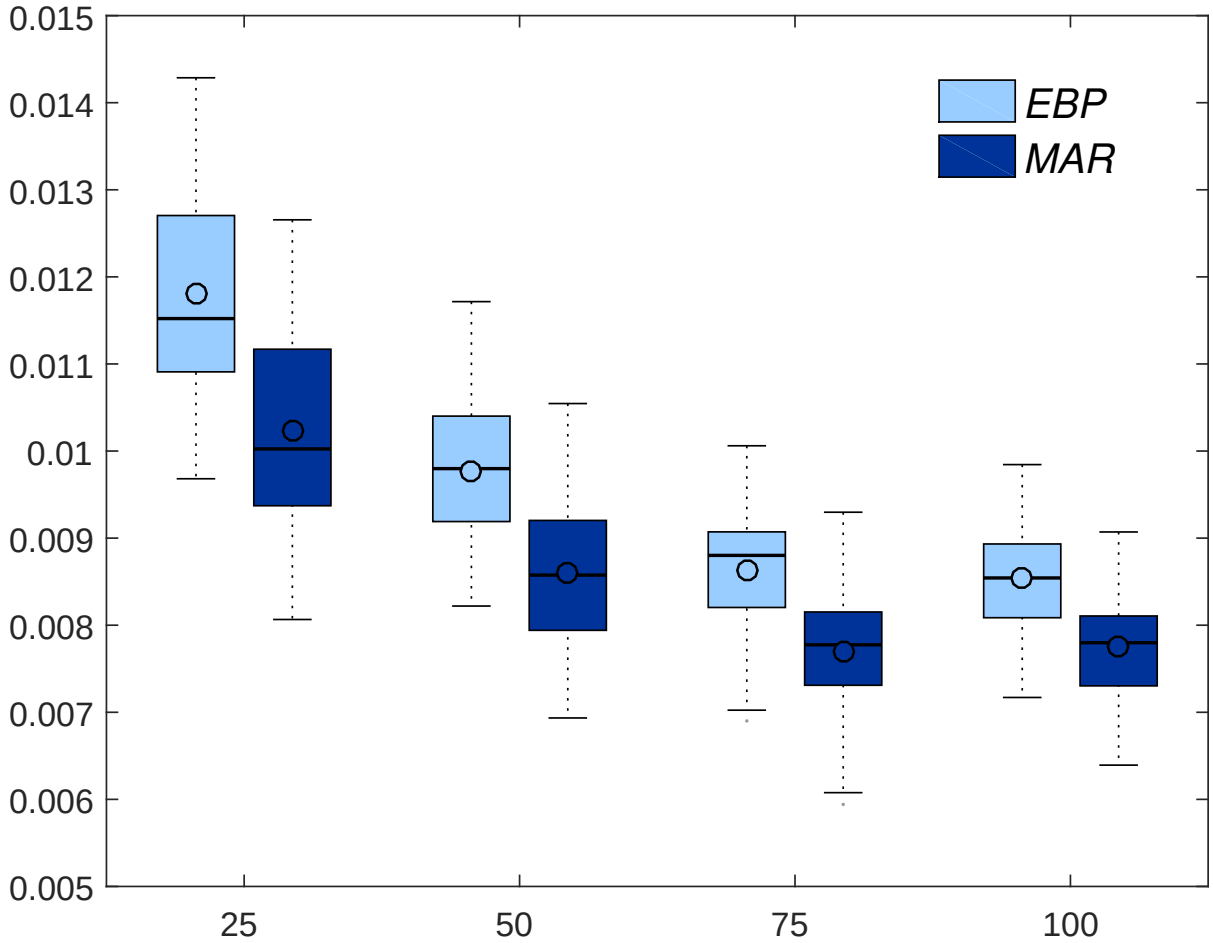

Figure 1. Boxplots of empirical biases B_d for EBP and MAR of proportions.

Simulation experiment

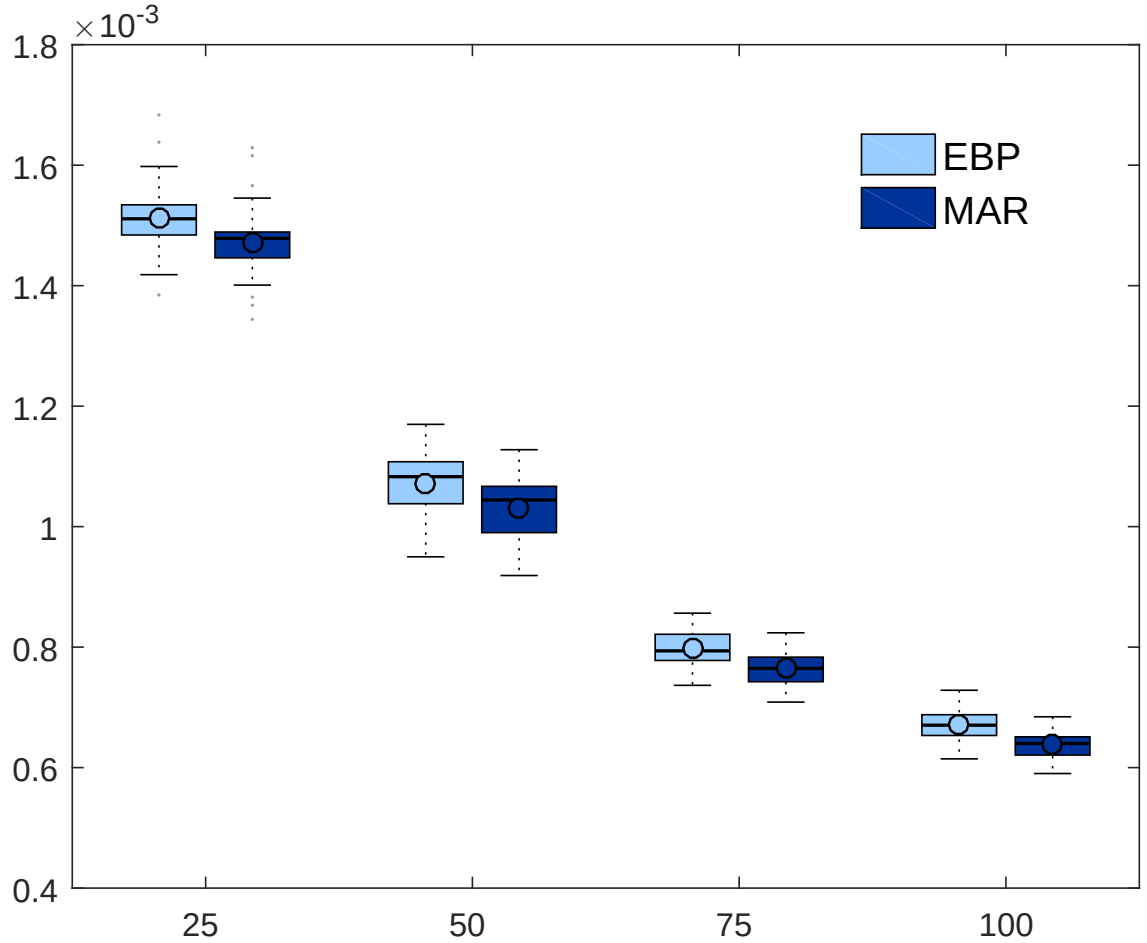

Figure 2. Boxplots of empirical MSEs E_d for EBP and MAR of proportions .

Simulation experiment

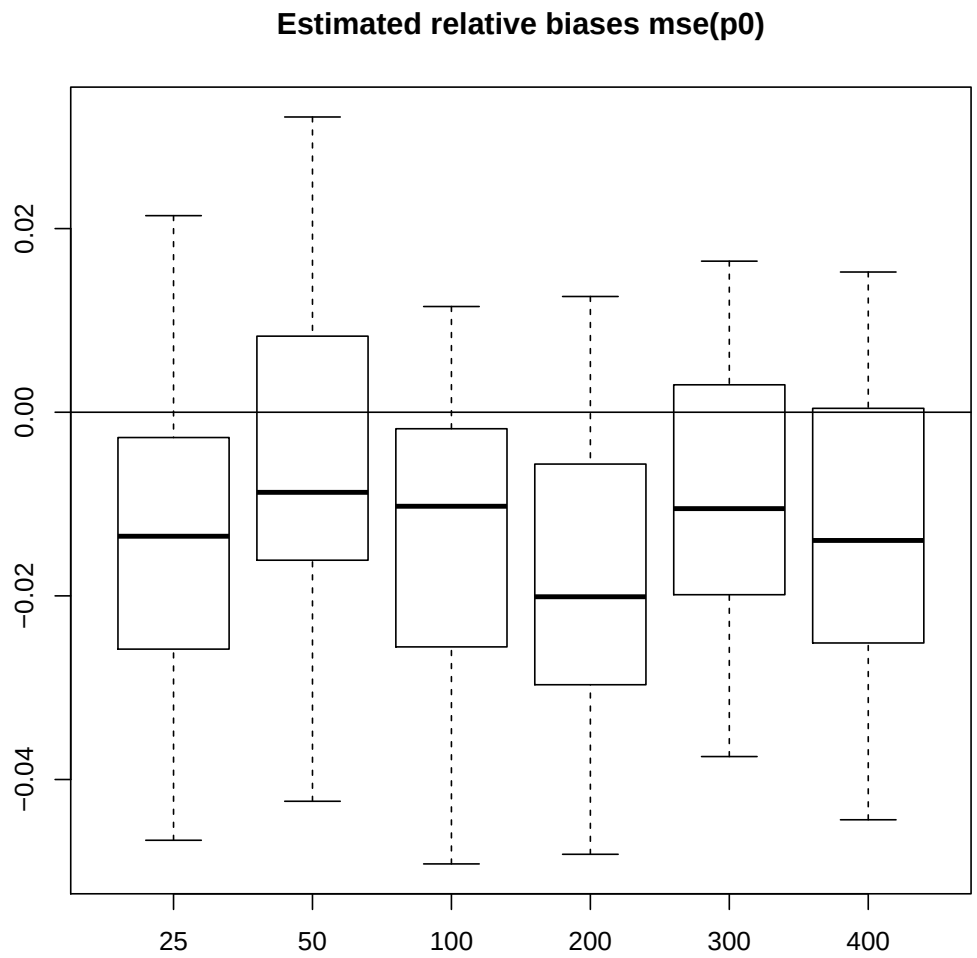

Figure 3. Relative biases of MSE estimators of MAR predictors for poverty proportions. Case $D = 30$, $n_d = 50$

Simulation experiment

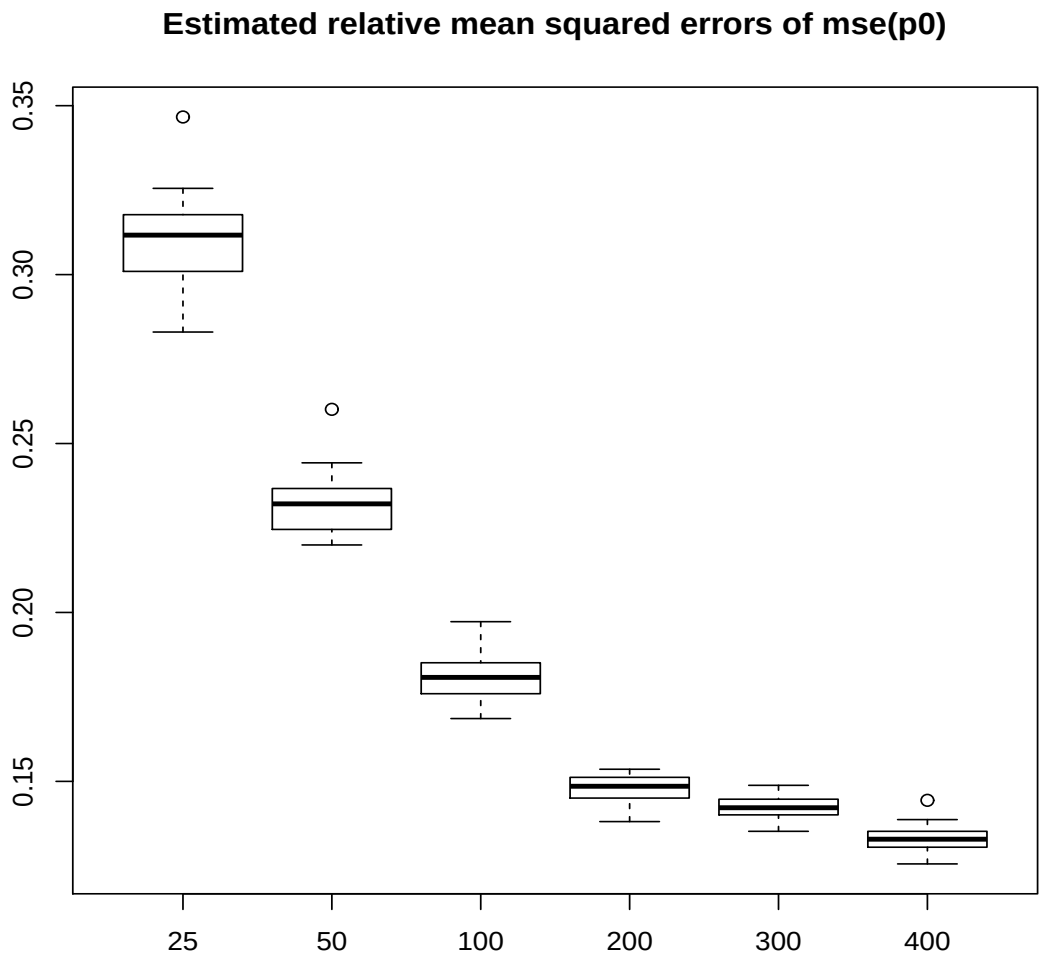

Figure 4. Relative root-MSEs of MSE estimators of MAR predictors for poverty proportions. Case $D = 30$, $n_d = 50$

Application to real data

Data from 2013 Spanish Living Conditions Survey (SLCS) in the Autonomous Community of Valencia

We are interested in estimating the domain poverty proportions in 2013

We consider $D = 26$ domains, comarcas (counties) appearing in the sample

Total sample size: $n = 2\,492$ *(SLCS 2013)*

Smallest area: 10 records

Largest area: 405 records

Population size: $N = 4\,877\,512$

Auxiliary agregated data (totals of covariate patterns) are taken from SLFS 2013

Application to real data

Data from 2013 Spanish Living Conditions Survey (SLCS) in the Autonomous Community of Valencia

We are interested in estimating the domain poverty proportions in 2013

We consider $D = 26$ domains, comarcas (counties) appearing in the sample

Total sample size: $n = 2\,492$ *(SLCS 2013)*

Smallest area: 10 records

Largest area: 405 records

Population size: $N = 4\,877\,512$

Auxiliary agregated data (totals of covariate patterns) are taken from SLFS 2013

Application to real data

- SLCS provides information regarding the **household income** received during the last year

- **Equivalent personal income**
  - is calculated in order to take into account scale economies in households
  - it is assigned to each member of the household (denoted as y_{dj}).

- **The poverty risk** is the proportion of people with equivalent personal income below the poverty threshold.

  E.g. the 2013 Valencia poverty threshold is $z = 6999.6$ (in EUR).

Application to real data

The model for personal income (in 10 000 EUR):

We assume that for $d = 1, \ldots, D,\ j = 1, \ldots, N_d$,

$$y_{dj}|_{v_d} \sim \text{Gamma}\big(\nu_{dj},\ a_{dj} = \frac{\nu_{dj}}{\mu_{dj}}\big),$$

where v_d are i.i.d. $N(0,1)$, $\nu_{dj} = a_{dj}\varphi$ and

$$g(\mu_{dj}) = \frac{1}{\mu_{dj}} = \beta_0 + \beta_1 \text{Employed}_{dj} + \beta_2 \text{Unemployed}_{dj} + \phi v_d$$

	estimate	standard error	p-value
β_0	0.775	0.0132	$<$ 2E-16
β_1	-0.141	0.0157	$<$ 2E-16
β_2	0.140	0.0300	3.09E-06
ϕ	0.1113	0.0112	$<$ 2E-16
φ	2.4646	0.0675	$<$ 2E-16

Table 1: Estimates of regression parameters.

Application to real data

The model for personal income (in 10 000 EUR):

We assume that for $d = 1, \ldots, D,\ j = 1, \ldots, N_d$,

$$y_{dj}|_{v_d} \sim \text{Gamma}\big(\nu_{dj},\ a_{dj} = \frac{\nu_{dj}}{\mu_{dj}}\big),$$

where v_d are i.i.d. $N(0,1)$, $\nu_{dj} = a_{dj}\varphi$ and

$$g(\mu_{dj}) = \frac{1}{\mu_{dj}} = \beta_0 + \beta_1 \text{Employed}_{dj} + \beta_2 \text{Unemployed}_{dj} + \phi v_d$$

	estimate	standard error	p-value
β_0	0.775	0.0132	$<$ 2E-16
β_1	-0.141	0.0157	$<$ 2E-16
β_2	0.140	0.0300	3.09E-06
ϕ	0.1113	0.0112	$<$ 2E-16
φ	2.4646	0.0675	$<$ 2E-16

Table 1: Estimates of regression parameters.

Application to real data

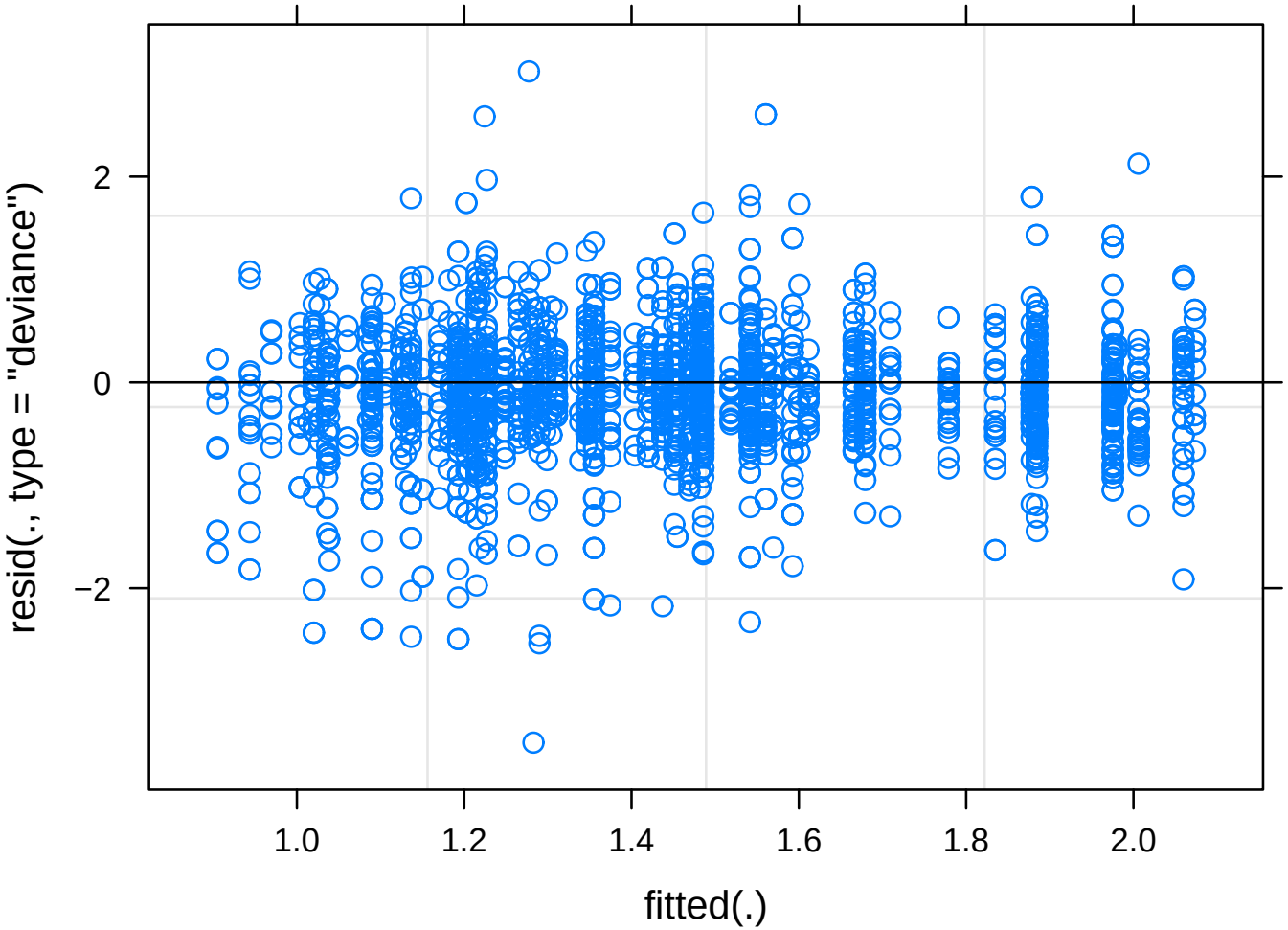

Figure 5. Plot of deviance residuals with respect to fitted values.

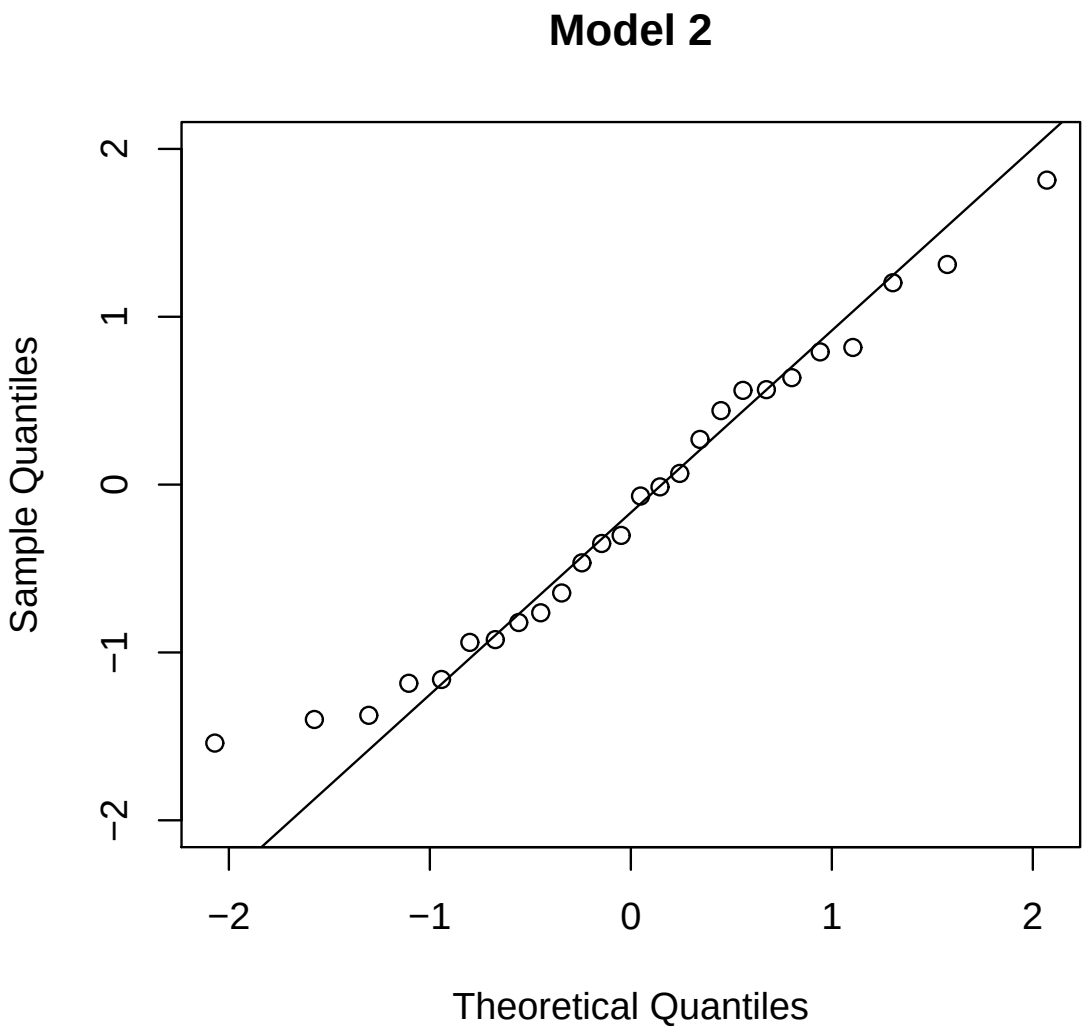

Figure 6. Q-Q plot of the predicted values of v_d.

Application to real data

Molina-Rao (CJS 2010) model:

- Let us consider the *log* transformation of data

$$z_{dj} = \log(y_{dj} + c)$$

and the nested error regression model

$$z_{dj} = \mathbf{x}_{dj}^T \boldsymbol{\beta} + u_d + e_{dj},$$

where $u_d \sim N(0, \sigma_u^2)$ and $e_{dj} \sim N(0, \sigma_e^2)$.

- $$r^2_{MolRao} = \sum_{d=1}^{D} \sum_{j=1}^{n_d} (y_{dj} - (\exp(\hat{z}_{dj}) - 1))^2 = 1938.30,$$

$$r^2_{Gamma} = \sum_{d=1}^{D} \sum_{j=1}^{n_d} (y_{dj} - \hat{\mu}_{dj})^2 = \mathbf{1897.35}.$$

Application to real data

Molina-Rao (CJS 2010) model:

- Let us consider the *log* transformation of data

$$z_{dj} = \log(y_{dj} + c)$$

and the nested error regression model

$$z_{dj} = \mathbf{x}_{dj}^T \boldsymbol{\beta} + u_d + e_{dj},$$

where $u_d \sim N(0, \sigma_u^2)$ and $e_{dj} \sim N(0, \sigma_e^2)$.

- $$r^2_{MolRao} = \sum_{d=1}^{D} \sum_{j=1}^{n_d} (y_{dj} - (\exp(\hat{z}_{dj}) - 1))^2 = 1938.30,$$

$$r^2_{Gamma} = \sum_{d=1}^{D} \sum_{j=1}^{n_d} (y_{dj} - \hat{\mu}_{dj})^2 = \mathbf{1897.35}.$$

Application to real data

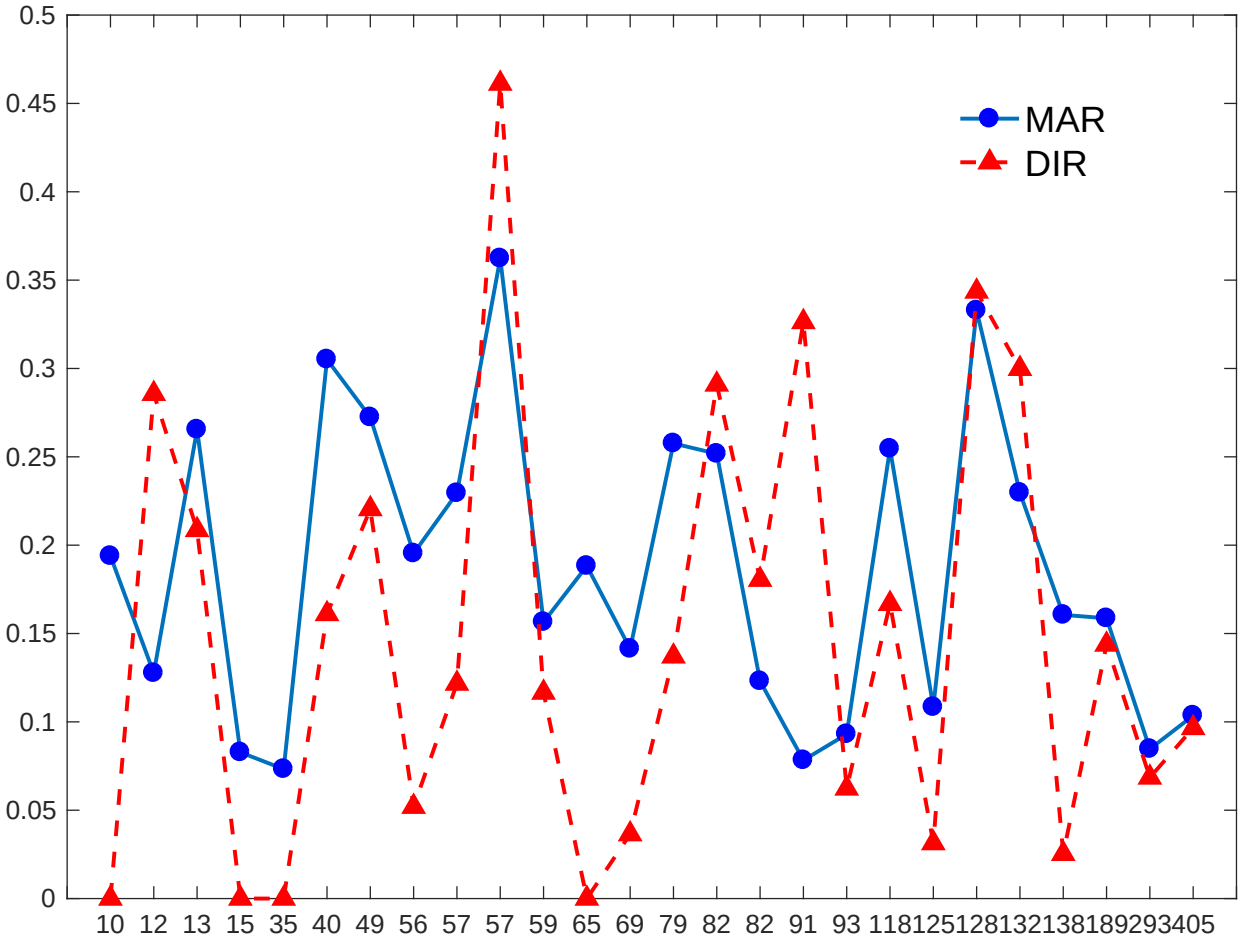

Figure 7. Marginal and Direct poverty proportions estimates.

Application to real data

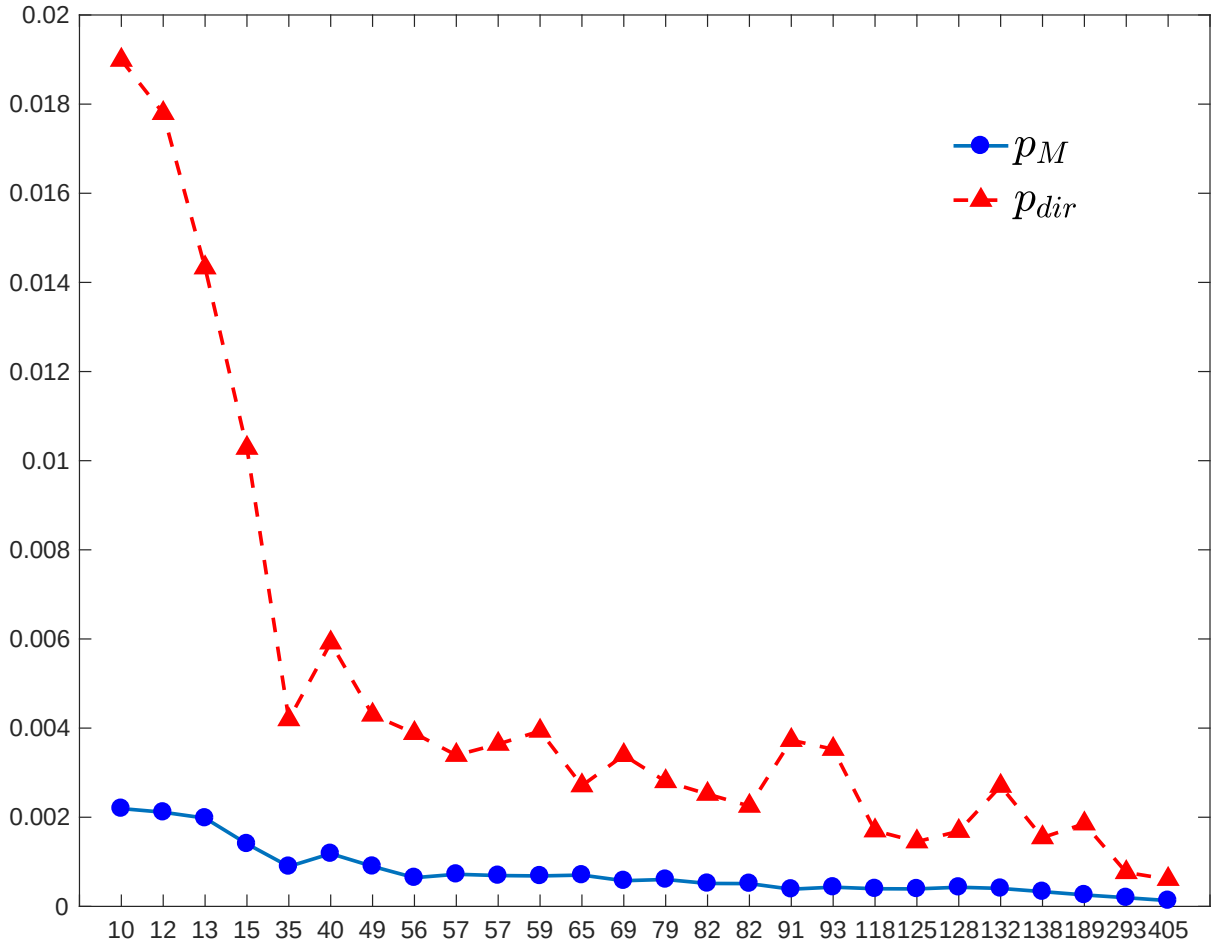

Figure 8. Estimated MSEs of poverty proportions estimates.

Conclusions:

- The proposed model and marginal predictor is applicable to small area estimation real data problems
- Marginal predictors can increase precision of the direct estimators

Thank you for your attention!!!