Multilevel Regression and Poststratification in Stata

Maurizio Pisati[1] Valeria Glorioso[1,2]

maurizio.pisati@unimib.it v.glorioso@campus.unimib.it

[1]Dept. of Sociology and Social Research
University of Milano-Bicocca (Italy)

[2]Dept. of Society, Human Development, and Health
Harvard School of Public Health

Stata Conference Chicago 2011
July 14-15

Outline

Outline

1. Introduction
   - The problem
   - The solution

2. Stata command

Outline

1. Introduction
   - The problem
   - The solution
2. Stata command
3. Simulations

Outline

1. Introduction
   - The problem
   - The solution
2. Stata command
3. Simulations
4. Conclusion

Outline

Introduction

A common research objective

- Sometimes social scientists are interested in determining whether, and to what extent, the distribution of a given target variable Y varies across K groups defined by the values of one or more covariates of interest

A common research objective

- Sometimes social scientists are interested in determining whether, and to what extent, the distribution of a given target variable Y varies across K groups defined by the values of one or more covariates of interest
- Let G denote a discrete variable representing the K groups under comparison. Without loss of generality, G can represent either a single discrete covariate or the cross-classification of two or more discrete covariates

A common research objective

- In symbols:

$$G = \prod_{g=1}^{M_G} V_g$$

where Π is the Cartesian product operator; M_G denotes the total number of covariates forming G; and V_g denotes the g^{th} covariate

A common research objective

- In symbols:

$$G = \prod_{g=1}^{M_G} V_g$$

where Π is the Cartesian product operator; M_G denotes the total number of covariates forming G; and V_g denotes the g^{th} covariate

- We will refer to G as the *group variable*

A common research objective

- The (conditional) distribution of Y within each category k of G can be described as follows:

$$Y_k \sim f(\theta_k, \phi_k) \quad \text{for } k = 1, \ldots, K$$

where $f(\cdot)$ denotes a generic probability distribution; θ_k denotes the expected value(s) of the distribution; and ϕ_k denotes one or more additional parameters of the distribution (e.g., its variance)

A common research objective

- For the sake of simplicity, let us focus on the expected value(s) of Y, so that our goal is to determine whether, and to what extent, the expected value(s) of Y varies/vary across the K categories of G

A common research objective

- For the sake of simplicity, let us focus on the expected value(s) of Y, so that our goal is to determine whether, and to what extent, the expected value(s) of Y varies/vary across the K categories of G
- In terms of regression analysis, this amounts to estimating the K possible values of the regression function $E(Y|G = k)$, i.e., $E(Y|G = 1) \equiv \theta_1$, $E(Y|G = 2) \equiv \theta_2$, ..., $E(Y|G = K) \equiv \theta_K$

A common research objective

- For the sake of simplicity, let us focus on the expected value(s) of Y, so that our goal is to determine whether, and to what extent, the expected value(s) of Y varies/vary across the K categories of G
- In terms of regression analysis, this amounts to estimating the K possible values of the regression function $E(Y|G=k)$, i.e., $E(Y|G=1) \equiv \theta_1$, $E(Y|G=2) \equiv \theta_2$, ..., $E(Y|G=K) \equiv \theta_K$
- Let us denote our estimand – i.e., our quantity of interest – by $\boldsymbol{\theta} \equiv \{\theta_k : k = 1, \ldots, K\}$

Estimating θ

- How do we get accurate – i.e., precise and unbiased – estimates of θ?

Estimating θ

- How do we get accurate – i.e., precise and unbiased – estimates of θ?
- For the sake of simplicity, let us suppose that (a) observations are sampled from a given target population, and (b) the data of interest are collected without measurement error, so that the only source of random estimation error is the sampling variance, and the only (possible) source of systematic estimation error is the selection bias

Estimating θ

- How do we get accurate – i.e., precise and unbiased – estimates of θ?
- For the sake of simplicity, let us suppose that (a) observations are sampled from a given target population, and (b) the data of interest are collected without measurement error, so that the only source of random estimation error is the sampling variance, and the only (possible) source of systematic estimation error is the selection bias
- The expression "selection bias" is used here as a shorthand for the sum of coverage bias, nonresponse bias, and sampling bias (Groves 1989)

Estimating $\boldsymbol{\theta}$

- The standard (maximum likelihood) estimator of each element θ_k of $\boldsymbol{\theta}$ is:

$$\hat{\theta}_k \equiv \widehat{E(Y|G = k)} = \frac{\sum_{i=1}^{n_k} Y_i}{n_k}$$

where n_k denotes the number of valid sample observations within category k of variable G

Estimating θ

- When n_k is small, $\hat{\theta}_k$ tends to be very unprecise, i.e., to generate highly variable estimates of θ_k

Estimating θ

- When n_k is small, $\hat{\theta}_k$ tends to be very unprecise, i.e., to generate highly variable estimates of θ_k
- The accuracy of $\hat{\theta}_k$ decreases further if the data object of analysis are affected by selection bias, i.e., if the valid observations are a nonrandom sample of the target population *and* the process of selection into the sample is associated with one or more variables that are also associated with variable Y

Here's Mr. P

- For all those cases where the number of valid observations within one or more categories of G is small and/or collected data are affected by selection bias, relatively accurate estimates of $\boldsymbol{\theta}$ can be obtained by using a proper combination of multilevel regression modeling and poststratification (henceforth MrP)

Here's Mr. P

- For all those cases where the number of valid observations within one or more categories of G is small and/or collected data are affected by selection bias, relatively accurate estimates of $\boldsymbol{\theta}$ can be obtained by using a proper combination of multilevel regression modeling and poststratification (henceforth MrP)
- This approach has been devised by Andrew Gelman and colleagues (Gelman and Little 1997; Park, Gelman and Bafumi 2004; Park, Gelman and Bafumi 2006; Gelman and Hill 2007) and recently elaborated on by Kastellec, Lax and Phillips (Lax and Phillips 2009a; Lax and Phillips 2009b; Kastellec, Lax and Phillips 2010)

The MrP estimator

- The MrP estimator of θ – which we will denote by $\tilde{\theta}$ – can be described as a four-step procedure as follows:

The MrP estimator

- **First:** Identify one or more covariates that might possibly be responsible for selection bias. Without loss of generality, let C denote a discrete variable representing the cross-classification of these covariates.

  In symbols:

$$C = \prod_{c=1}^{M_C} V_c$$

  where Π is the Cartesian product operator; M_C denotes the total number of covariates forming C; and V_c denotes the c^{th} covariate.

  We will refer to C as the *composition variable*

The MrP estimator

- **Second:** Define the new estimand $\boldsymbol{\gamma} \equiv \{\gamma_{kl} : k = 1, \dots, K; l = 1, \dots, L\}$, where $\gamma_{kl} \equiv E(Y|G = k, C = l)$; k indexes the K categories of variable G as above; and l indexes the L categories of variable C

The MrP estimator

- **Third:** Use a properly specified multilevel regression model to estimate $\boldsymbol{\gamma}$

The MrP estimator

- **Fourth:** Compute the estimate of each element θ_k of $\boldsymbol{\theta}$ as a weighted sum of the proper subset of $\hat{\boldsymbol{\gamma}}$:

$$\tilde{\theta}_k = \sum_{l=1}^{L} \hat{\gamma}_{kl} w_{l|k}$$

where $w_{l|k} = N_{kl}/N_k$; N_k denotes the number of members of the target population who belong in category k of variable G; and N_{kl} denotes the number of members of the target population who belong in category k of variable G *and* in category l of variable C

The MrP estimator: Advantages

- The use of multilevel regression modeling (step 3 above) helps to increase precision

The MrP estimator: Advantages

- The use of multilevel regression modeling (step 3 above) helps to increase precision
- If the composition variable C is carefully defined, poststratification (step 4 above) helps to decrease bias

The MrP estimator: Advantages

- The use of multilevel regression modeling (step 3 above) helps to increase precision
- If the composition variable C is carefully defined, poststratification (step 4 above) helps to decrease bias
- In sum, we expect MrP to be a relatively accurate estimator of θ

The MrP estimator: Disadvantages

- We need to have population data – or, at least, a sufficiently accurate estimate of it – for the full $G \times C$ cross-classification; this might limit the definition of C

The MrP estimator: Disadvantages

- We need to have population data – or, at least, a sufficiently accurate estimate of it – for the full $G \times C$ cross-classification; this might limit the definition of C
- To get good estimates of $\boldsymbol{\gamma}$, the multilevel regression model must be specified very carefully – but this caveat applies to any kind of regression model

Stata command

mrp – a Stata implementation of MrP

- mrp is a novel user-written Stata command that implements the MrP estimator outlined above

mrp – a Stata implementation of MrP

- mrp is a novel user-written Stata command that implements the MrP estimator outlined above
- Basically, mrp requests the user to specify (a) the target variable Y; (b) the list of covariates forming the group variable G; (c) the list of covariates forming the composition variable C; (d) the multilevel regression command appropriate to the problem at hand (e.g., xtmixed); (e) the list of "fixed effects"; (f) the list of "random effects"; and (g) the name of a properly arranged dataset contaning the population totals N_{kl}

mrp – a Stata implementation of MrP

- mrp is a novel user-written Stata command that implements the MrP estimator outlined above
- Basically, mrp requests the user to specify (a) the target variable Y; (b) the list of covariates forming the group variable G; (c) the list of covariates forming the composition variable C; (d) the multilevel regression command appropriate to the problem at hand (e.g., xtmixed); (e) the list of "fixed effects"; (f) the list of "random effects"; and (g) the name of a properly arranged dataset containing the population totals N_{kl}
- The basic output of mrp is an estimate of the K values of the regression function $E(Y|G = k)$, i.e., of the K elements of $\boldsymbol{\theta}$

Example (based on simulation)

- Our objective is to describe the extent to which the proportion of Italian adults who attend Catholic Mass regularly varies across Italian regions

Example (based on simulation)

- Our objective is to describe the extent to which the proportion of Italian adults who attend Catholic Mass regularly varies across Italian regions
- To this aim, a simple random sample of 2,000 units is drawn from the target population (Italian men and women aged 18+), and each sampled unit is contacted for interview

Example (based on simulation)

- Our objective is to describe the extent to which the proportion of Italian adults who attend Catholic Mass regularly varies across Italian regions
- To this aim, a simple random sample of 2,000 units is drawn from the target population (Italian men and women aged 18+), and each sampled unit is contacted for interview
- Only 984 subjects accept to participate in the survey. The response rate turns out to be higher among women and positively correlated with age and educational level

Example

- Since the number of valid observations within each region k is generally small ($\min(n_k)$=30, $\max(n_k)$=97), the standard estimator of $\boldsymbol{\theta}$ will be very unprecise

Example

- Since the number of valid observations within each region k is generally small ($\min(n_k)=30$, $\max(n_k)=97$), the standard estimator of θ will be very unprecise
- Moreover, since sex, age, and educational level are associated with Catholic Mass attendance, the standard estimator of θ will likely be affected by selection bias

Example

- Since the number of valid observations within each region k is generally small (min(n_k)=30, max(n_k)=97), the standard estimator of θ will be very unprecise
- Moreover, since sex, age, and educational level are associated with Catholic Mass attendance, the standard estimator of θ will likely be affected by selection bias
- In an attempt to increase precision and decrease bias, we estimate θ using the new Stata command `mrp`

Example: mrp specification

Target variable

```
mrp church, g(region relmar|region) c(sex age edu)   ///
    regcommand(xtmixed) binomial                     ///
    fe(relmar) re(i.age i.edu i.sex i.region)        ///
    popref("PopRef.dta") npop(N)                     ///
    percent
```

Example: mrp specification

List of covariates forming group variable G

```
mrp church, g(region relmar|region) c(sex age edu)   ///
    regcommand(xtmixed) binomial                      ///
    fe(relmar) re(i.age i.edu i.sex i.region)         ///
    popref("PopRef.dta") npop(N)                      ///
    percent
```

Example: mrp specification

List of covariates forming composition variable C

```
mrp church, g(region relmar|region) c(sex age edu)    ///
    regcommand(xtmixed) binomial                      ///
    fe(relmar) re(i.age i.edu i.sex i.region)         ///
    popref("PopRef.dta") npop(N)                      ///
    percent
```

Example: mrp specification

Multilevel regression command

```
mrp church, g(region relmar|region) c(sex age edu)   ///
    regcommand(xtmixed) binomial                      ///
    fe(relmar) re(i.age i.edu i.sex i.region)         ///
    popref("PopRef.dta") npop(N)                      ///
    percent
```

Example: mrp specification

List of “fixed effects”

```
mrp church, g(region relmar|region) c(sex age edu)   ///
    regcommand(xtmixed) binomial                     ///
    fe(relmar) re(i.age i.edu i.sex i.region)        ///
    popref("PopRef.dta") npop(N)                     ///
    percent
```

Example: `mrp` specification

List of "random effects"

```
mrp church, g(region relmar|region) c(sex age edu)   ///
    regcommand(xtmixed) binomial                     ///
    fe(relmar) re(i.age i.edu i.sex i.region)        ///
    popref("PopRef.dta") npop(N)                     ///
    percent
```

Example: mrp specification

Dataset and variable containing population totals N_{kl}

```
mrp church, g(region relmar|region) c(sex age edu)   ///
    regcommand(xtmixed) binomial                      ///
    fe(relmar) re(i.age i.edu i.sex i.region)         ///
    popref("PopRef.dta") npop(N)                      ///
    percent
```

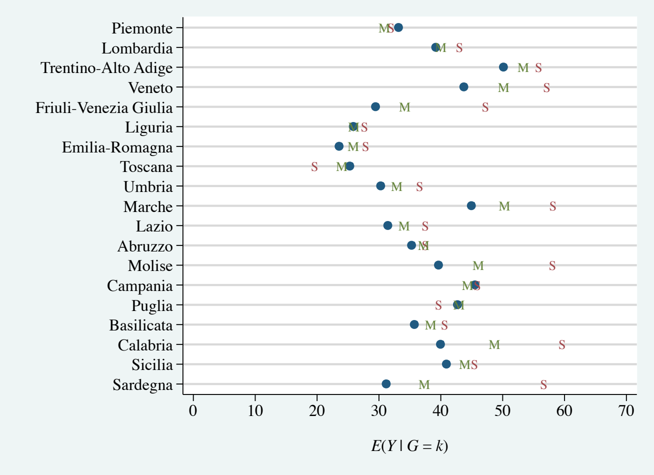
Example: mrp specification

Scale option (converts proportions into percentages)

```
mrp church, g(region relmar|region) c(sex age edu)   ///
    regcommand(xtmixed) binomial                     ///
    fe(relmar) re(i.age i.edu i.sex i.region)        ///
    popref("PopRef.dta") npop(N)                     ///
    percent
```

Example: Results

Dot = True population value, S = Standard estimate, M = MrP estimate



Simulations

Quantity of interest

- We used Monte Carlo simulation to evaluate the performance of three estimators of θ in a research setting analogous to the one illustrated in the example above

Quantity of interest

- We used Monte Carlo simulation to evaluate the performance of three estimators of θ in a research setting analogous to the one illustrated in the example above
- Our underlying research objective is to describe the extent to which the proportion of Italian adults who attend Catholic Mass regularly varies across 19 of the 20 regions into which Italy is subdivided (the 20^{th} region, Valle d'Aosta, is excluded from the analysis because of its peculiarities)

Quantity of interest

- We used Monte Carlo simulation to evaluate the performance of three estimators of θ in a research setting analogous to the one illustrated in the example above
- Our underlying research objective is to describe the extent to which the proportion of Italian adults who attend Catholic Mass regularly varies across 19 of the 20 regions into which Italy is subdivided (the 20^{th} region, Valle d'Aosta, is excluded from the analysis because of its peculiarities)
- Thus, our quantity of interest θ corresponds to the $K = 19$ values of the regression function $E(Y|G = k)$, where the target variable Y is a binary indicator of regular Catholic Mass attendance, and the group variable G is the region of residence

Procedure

- For each estimator, we followed a three-step procedure:

Procedure

- For each estimator, we followed a three-step procedure:
  1. First, we simulated 1,000 sample surveys, using as the sampling frame a large dataset ($N = 251,708$) that mimics the socio-demographic structure of the full Italian adult population and contains complete information on the following individual characteristics: region of residence (`region`), sex (`sex`), age (`age`), educational level (`edu`), and Catholic Mass attendance (`church`)

Procedure

- For each estimator, we followed a three-step procedure:
  1. First, we simulated 1,000 sample surveys, using as the sampling frame a large dataset ($N = 251,708$) that mimics the socio-demographic structure of the full Italian adult population and contains complete information on the following individual characteristics: region of residence (`region`), sex (`sex`), age (`age`), educational level (`edu`), and Catholic Mass attendance (`church`)
  2. Second, we used the data collected in each simulated survey to estimate the quantity of interest, thus getting a simulated sampling distribution of θ made of 1,000 estimates

Procedure

- For each estimator, we followed a three-step procedure:
  1. First, we simulated 1,000 sample surveys, using as the sampling frame a large dataset ($N = 251,708$) that mimics the socio-demographic structure of the full Italian adult population and contains complete information on the following individual characteristics: region of residence (`region`), sex (`sex`), age (`age`), educational level (`edu`), and Catholic Mass attendance (`church`)
  2. Second, we used the data collected in each simulated survey to estimate the quantity of interest, thus getting a simulated sampling distribution of θ made of 1,000 estimates
  3. Finally, we evaluated the estimator in question by computing its bias, empirical standard error, and root mean square error

Survey specifications

- Sampling method: Simple random sampling

Survey specifications

- Sampling method: Simple random sampling
- Initial sample size: $n = 2,000$

Survey specifications

- Sampling method: Simple random sampling
- Initial sample size: $n = 2,000$
- Response rate: Each sampled unit is selected into the final sample with a probability determined by his/her sex, age, and educational level. Such probabilities range from a minimum of 20% (poorly-educated men aged 18-44) to a maximum of 100 % (highly-educated women aged 65+)

Survey specifications

- Sampling method: Simple random sampling
- Initial sample size: $n = 2,000$
- Response rate: Each sampled unit is selected into the final sample with a probability determined by his/her sex, age, and educational level. Such probabilities range from a minimum of 20% (poorly-educated men aged 18-44) to a maximum of 100 % (highly-educated women aged 65+)
- Final sample size: mean(n)=970, min(n)=897, max(n)=1,035

Estimator 1

Standard (`std`)

- The standard estimator of each element θ_k of $\boldsymbol{\theta}$ is defined as follows:

$$\hat{\theta}_k = \frac{\sum_{i=1}^{n_k} \texttt{church}_i}{n_k}$$

where $\texttt{church}_i$ takes value 1 when subject i attends Catholic Mass regularly, value 0 otherwise; and n_k denotes the number of valid sample observations within region of residence k

Estimator 2

Multilevel Regression with Poststratification (`mrp`)

- The MrP estimator of each element θ_k of $\boldsymbol{\theta}$ is defined as follows:

$$\tilde{\theta}_k = \sum_{l=1}^{L} \hat{\gamma}_{kl} w_{l|k}$$

where all symbols are defined as in slides 14-16 above

Estimator 2

Multilevel Regression with Poststratification (`mrp`)

- The MrP estimator of each element θ_k of $\boldsymbol{\theta}$ is defined as follows:

$$\tilde{\theta}_k = \sum_{l=1}^{L} \hat{\gamma}_{kl} w_{l|k}$$

where all symbols are defined as in slides 14-16 above

- The estimation of parameters γ_{kl} requires that the composition variable C be previously defined

Estimator 2

Multilevel Regression with Poststratification (`mrp`)

- The MrP estimator of each element θ_k of $\boldsymbol{\theta}$ is defined as follows:

$$\tilde{\theta}_k = \sum_{l=1}^{L} \hat{\gamma}_{kl} w_{l|k}$$

  where all symbols are defined as in slides 14-16 above
- The estimation of parameters γ_{kl} requires that the composition variable C be previously defined
- In our case, we define C as the cross-classification of three categorical covariates: `sex` (2 levels), `age` (4 levels), and `edu` (3 levels). Therefore, $L = 2 \times 4 \times 3 = 24$

Estimator 2

Multilevel Regression with Poststratification (**mrp**)

- Given the definition of composition variable C, the parameters γ_{kl} are estimated using the following multilevel regression model:

$$\gamma_{kl} = \beta_0 + \alpha_k^{\texttt{region}} + \alpha_{r[l]}^{\texttt{sex}} + \alpha_{s[l]}^{\texttt{age}} + \alpha_{t[l]}^{\texttt{edu}}$$

Estimator 2

Multilevel Regression with Poststratification (mrp)

where

$$\begin{aligned}
\alpha_k^{\texttt{region}} &\sim N(\beta^{\texttt{relmar}} \cdot \texttt{relmar}, \sigma^2_{\texttt{region}}) \quad \text{for } k = 1, \ldots, 19 \\
\alpha_r^{\texttt{sex}} &\sim N(0, \sigma^2_{\texttt{sex}}) \quad \text{for } r = 1, \ldots, 2 \\
\alpha_s^{\texttt{age}} &\sim N(0, \sigma^2_{\texttt{age}}) \quad \text{for } s = 1, \ldots, 4 \\
\alpha_t^{\texttt{edu}} &\sim N(0, \sigma^2_{\texttt{edu}}) \quad \text{for } t = 1, \ldots, 3
\end{aligned}$$

and `relmar` is a region-level variable that expresses the percentage of religious marriages in each region

Estimator 3

Standard Regression with Poststratification (**srp**)

- The SrP estimator of each element θ_k of $\boldsymbol{\theta}$ is defined as follows:

$$\ddot{\theta}_k = \sum_{l=1}^{L} \hat{\gamma}_{kl} w_{l|k}$$

where all symbols are defined as above

Estimator 3

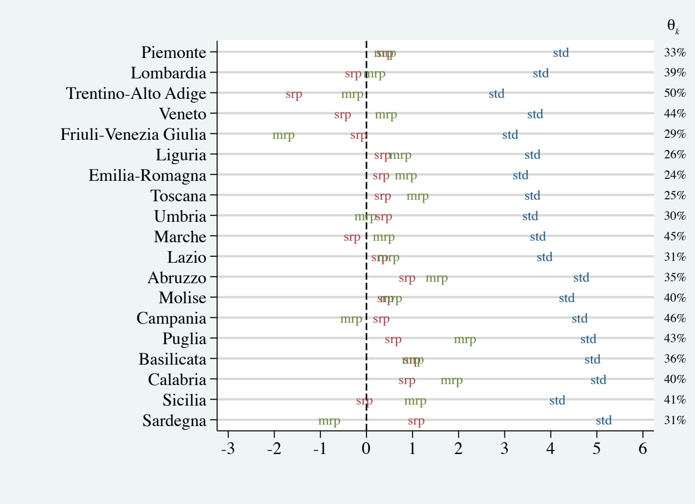
Standard Regression with Poststratification (**srp**)

- The SrP estimator has the same general form as the MrP estimator, but in the SrP estimator the parameters γ_{kl} are estimated using a standard logistic regression model as follows:

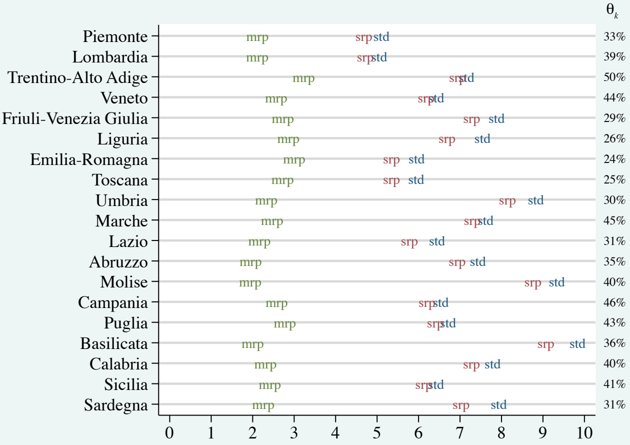
$$\gamma_{kl} = \text{invlogit}(\beta_0 + \beta^{\texttt{relmar}} \cdot \texttt{relmar} + \beta_k^{\texttt{region}} \cdot \texttt{region}_k + \\ + \beta_r^{\texttt{sex}} \cdot \texttt{sex}_{r[l]} + \beta_s^{\texttt{age}} \cdot \texttt{age}_{s[l]} + \beta_t^{\texttt{edu}} \cdot \texttt{edu}_{t[l]})$$

where $\beta_1^{\texttt{region}} = \beta_2^{\texttt{region}} = \beta_1^{\texttt{sex}} = \beta_1^{\texttt{age}} = \beta_1^{\texttt{edu}} = 0$

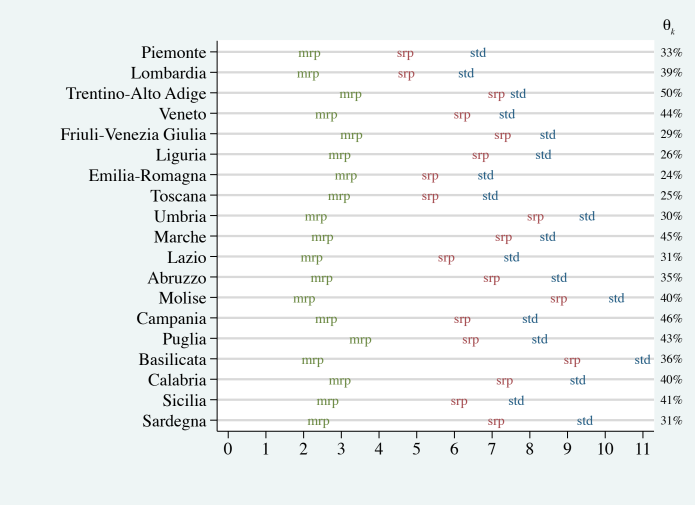
Results: Bias



Results: Empirical standard error



Results: Root mean square error



Results: Summary

- **Bias**: In absolute terms, the MrP estimator exhibits little bias – in most cases less than one percentage point, for an average true value of 36%. Comparatively, it exhibits significantly less bias than the standard estimator and slightly more bias than the SrP estimator

Results: Summary

- **Bias**: In absolute terms, the MrP estimator exhibits little bias – in most cases less than one percentage point, for an average true value of 36%. Comparatively, it exhibits significantly less bias than the standard estimator and slightly more bias than the SrP estimator
- **Precision**: The MrP estimator is significantly more precise (i.e., less variable) than both the standard estimator and the SrP estimator

Results: Summary

- **Bias**: In absolute terms, the MrP estimator exhibits little bias – in most cases less than one percentage point, for an average true value of 36%. Comparatively, it exhibits significantly less bias than the standard estimator and slightly more bias than the SrP estimator
- **Precision**: The MrP estimator is significantly more precise (i.e., less variable) than both the standard estimator and the SrP estimator
- **Accuracy**: Combining bias and precision, we can conclude that the MrP estimator is 1 to 4 times more accurate than the standard estimator and 1 to 3 times more accurate than the SrP estimator

Conclusion

Conclusion

- `mrp` is still at alpha stage and it will take a few months before it reaches a publishable form

Conclusion

- `mrp` is still at alpha stage and it will take a few months before it reaches a publishable form
- `mrp` is part of a larger project on the analysis of variation in Stata, and eventually it will be subsumed under a more general command for the analysis of association

Conclusion

- `mrp` is still at alpha stage and it will take a few months before it reaches a publishable form
- `mrp` is part of a larger project on the analysis of variation in Stata, and eventually it will be subsumed under a more general command for the analysis of association
- Part of the work presented here was carried out while Maurizio Pisati was a visiting scholar at the Institute for Quantitative Social Science at Harvard University, and Valeria Glorioso was a visiting student researcher at the Department of Society, Human Development, and Health of the Harvard School of Public Health

References

References

- Gelman, A. and J. Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Gelman, A. and T.C. Little. 1997. Poststratification into many categories using hierarchical logistic regression. *Survey Methodology* 23: 127–135.
- Groves, R.M. 1989. *Survey Errors and Survey Costs*. New York: Wiley.
- Kastellec, J., Lax, J.R. and J.H. Phillips. 2010. Public opinion and Senate confirmation of Supreme Court nominees. *Journal of Politics* 72: 767–784.
- Lax, J.R. and J.H. Phillips. 2009a. How should we estimate public opinion in the States?. *American Journal of Political Science* 53: 107–121.
- Lax, J.R. and J.H. Phillips. 2009b. Gay rights in the States: Public opinion and policy responsiveness. *American Political Science Review* 103: 367–386.
- Park, D.K., Gelman, A. and J. Bafumi. 2004. Bayesian multilevel estimation with poststratification: State-level estimates from national polls. *Political Analysis* 12: 375–385.
- Park, D.K., Gelman, A. and J. Bafumi. 2006. State level opinions from national surveys: Poststratification using multilevel logistic regression. In *Public Opinion in State Politics*. Ed. J.E. Cohen. Stanford, CA: Stanford University Press, 209–228.