



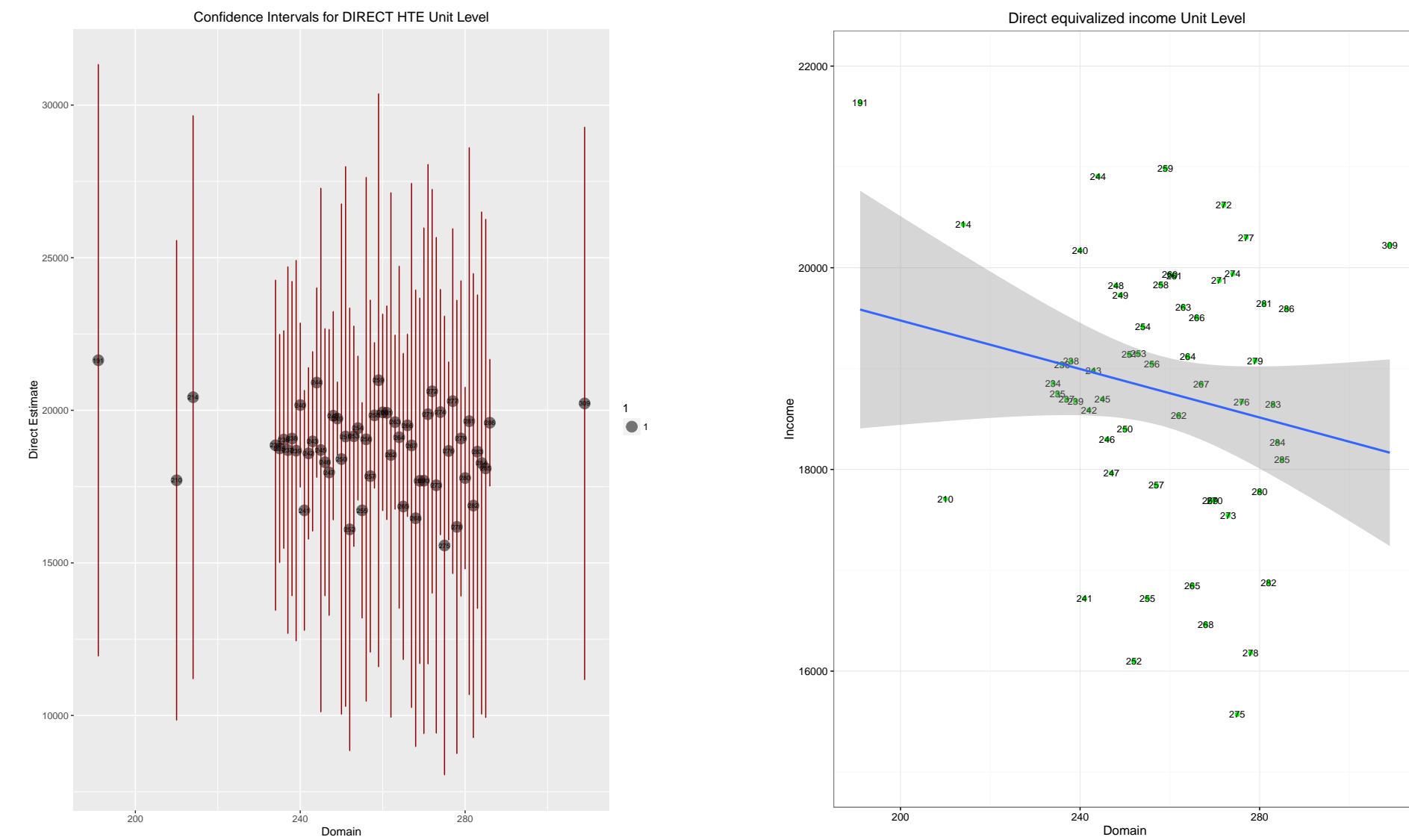
ESTIMATING INCOME OF TUSCAN HOUSEHOLDS

TOMMASO CRESCIOLI, ANNA MERGONI & MARCO RISPOLI



PROBLEM

In order to implement correct policies to fight poverty we need to have data at NUTS III level. Usually at this level few data are available, so direct estimates are not reliable due to the wide CV they have:



Following the National Institute of Statistics, estimates with coefficient of variation > 20% are not reliable, therefore we need to switch to SAE methods.

CONTRIBUTION

What did we do?

1. Calculate SAE methods: *EBLUP unit level*, *FH* and *GREG*.
2. Comparison of the results with direct estimates in order to asses if there is an improvement.
3. The choice of *the best model* given the characteristics of our data and covariates.

DATA

We used data unit level on Equivalized Income on LLS and other 5 covariates (m^2 House, Education, Gender, Single, Work) from EUSILC.

- Survey A: 1426 households (2008)
- Census: 1388252 households (2001)

**only the significant covariates were selected*

METHODOLOGY

1. *Direct*: for large enough sample size.

$$\hat{Y}_d^{DIR} = N_d^{-1} * \sum_{j \in s_d} w_{dj} * Y_{dj}$$

$$d = 1, \dots, D$$

2. *Eblup unit level*: used when auxiliary information available at unit.

$$\hat{\theta}_d^E = \gamma_d [\bar{y}_d + (\bar{X}_d^T - \bar{x}_d^T) \hat{\beta}] + (1 - \gamma_d) \bar{x}_d^T \hat{\beta}$$

$$\gamma_d = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \hat{\sigma}_e^2 / n_d} \quad d = 1, \dots, D$$

3. *GREG*: model assisted estimator.

$$\hat{t}_G = \sum_{j \in s_d} w_{dj} y_{dj} + \hat{\beta}^T (t_x - \sum_{j \in s} w_{dj} x_{dj})$$

$$d = 1, \dots, D \quad j = 1, \dots, N_d$$

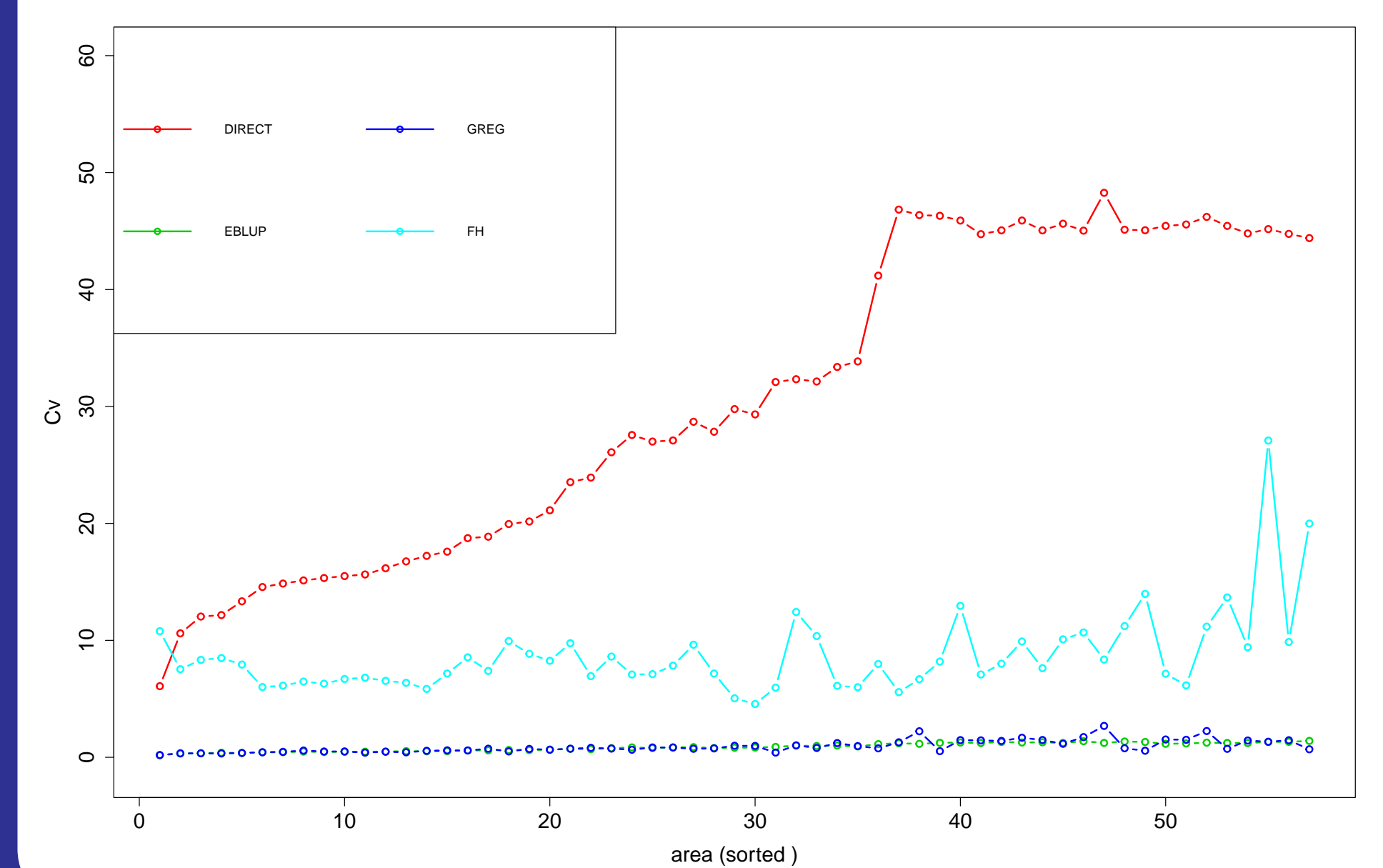
4. *Fay - Herriot*: two stage model based.

$$\hat{\delta}_d^{DIR} = x_d^T * \beta + v_d + e_d \quad d = 1, \dots, D$$

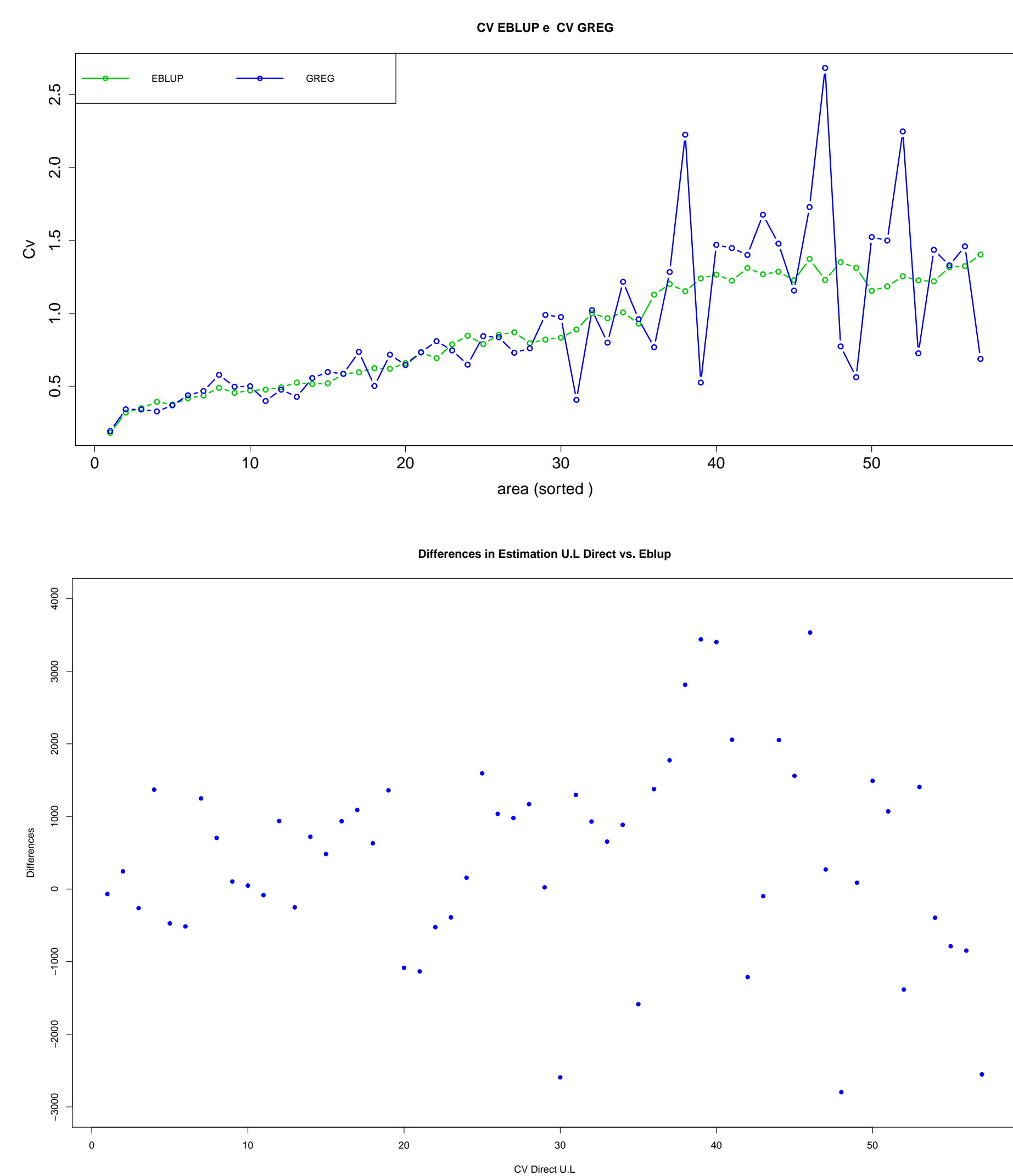
RESULTS

LLS	n_i	DIRECT	EBLUP	GREG	FH	CV_{dir}	CV_{eblup}	CV_{greg}	CV_{FH}
249	281	19730	19798	19800	19707	6.08	0.18	0.192	10.784
286	92	19592	19348	19343	19256	10.6	0.319	0.341	7.534
242	45	18588	17884	17837	18523	15.13	0.489	0.578	6.474
274	25	19939	18580	18563	18472	20.16	0.62	0.716	8.848
237	10	18695	17765	17745	17768	32.14	0.966	0.798	10.365
309	5	20224	16785	16534	16432	44.76	1.323	1.459	9.863

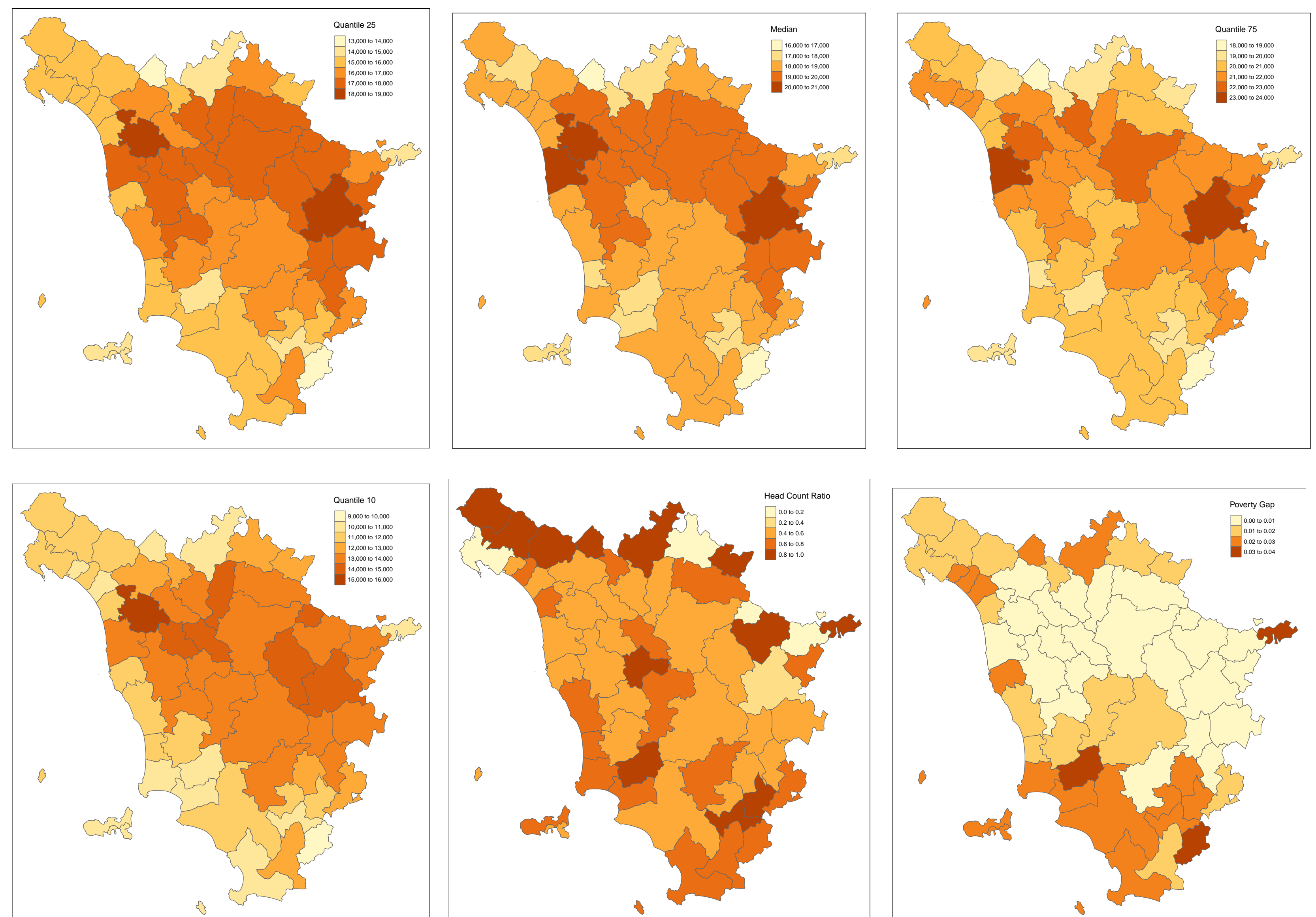
COMPARING CV



A WINDOW ON UNIT LEVEL



POVERTY MAPPING



IMPLEMENTATION

We implement our methodologies using R, specifically :

```
library(sae)
library(ggplot2)
library(emdi)
library(JoSAE)
```

REFERENCES

- [1] EUROSTAT: (2014) Handbook on Methodology of Modern Business Statistics In CROS.
- [2] I. Molina and Y. Marhuenda: (2015) An R Package for Small Area Estimation. In *The R Journal* Vol. 7/1, June
- [3] M. Pratesi: (2016) Analysis of Poverty Data by Small Area Estimation. In *Wiley Series in Survey Methodology*

ACKNOWLEDGEMENT

We would like to acknowledge Prof. Monica Pratesi, for the opportunities she gave us, Research Fellow Gaia Bertarelli, for the evergreen support and Prof. Stefano Marchetti for his precious hints