



# ESTIMATING EQUIVALIZED INCOME IN TUSCANY WITH SAE METHODS

Tural Ahmadov , Merve Celimli - University of Pisa, MSc in Economics

## ABSTRACT

Small area estimation is becoming important in survey sampling due to a growing demand for reliable Small Area Statistics from both public and private sectors. It is now widely recognized that Direct Survey Estimates for small areas are likely to yield unacceptably large standard errors due to the smallness of sample sizes in the areas. This makes it necessary to "borrow strength" from related areas to find more accurate estimates for a given area or, simultaneously, for several areas. This has led to the development of alternative methods such as synthetic, sample size dependent, empirical best linear unbiased prediction, etc. The present poster is largely an appraisal of some of these methods.

## CONTACT:

Tural Ahmadov  
University of Pisa  
Email: tural.ahmadov095@gmail.com

Merve Celimli  
University of Pisa  
Email: mervecelimli09@gmail.com

## DATA INTRODUCTION

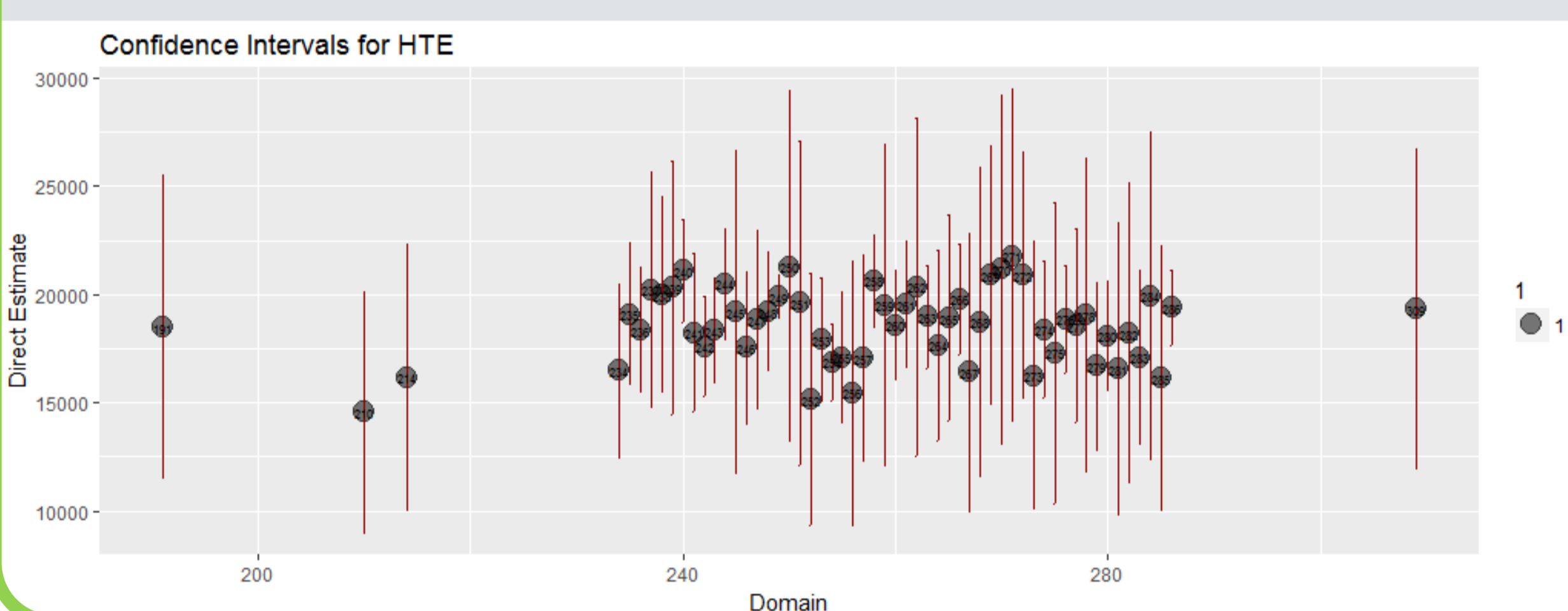
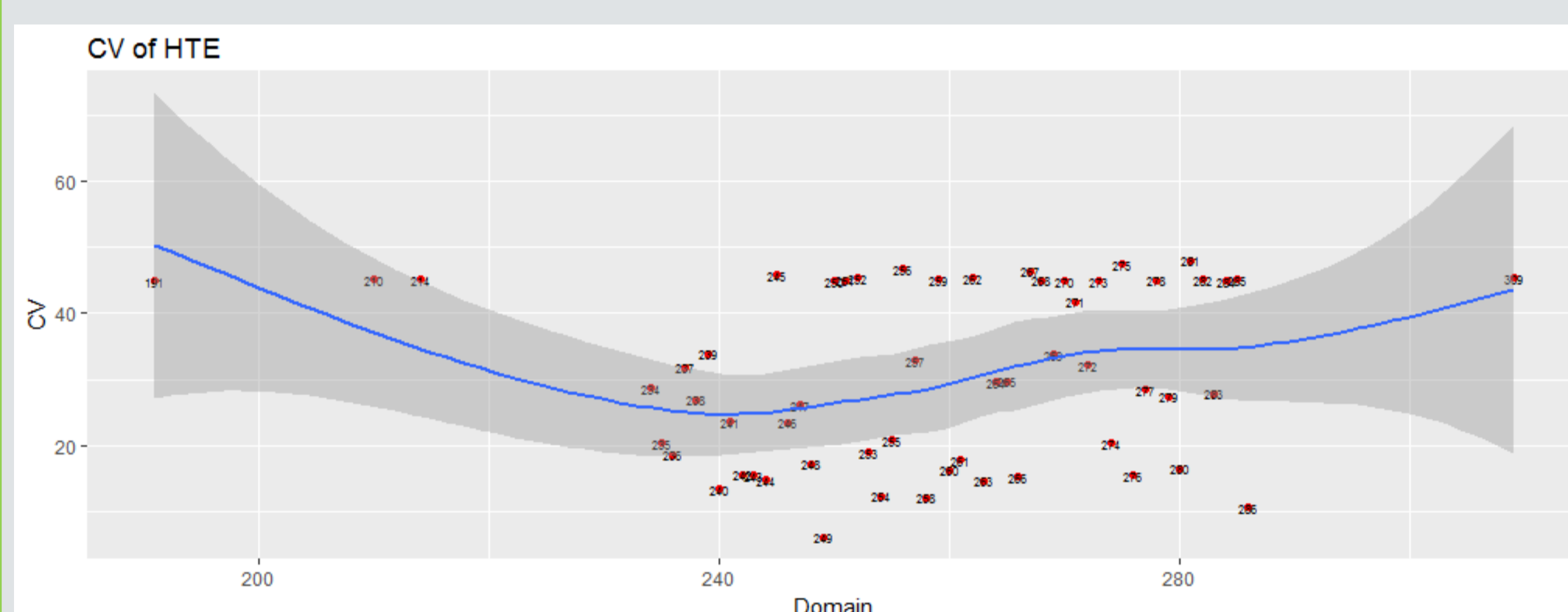
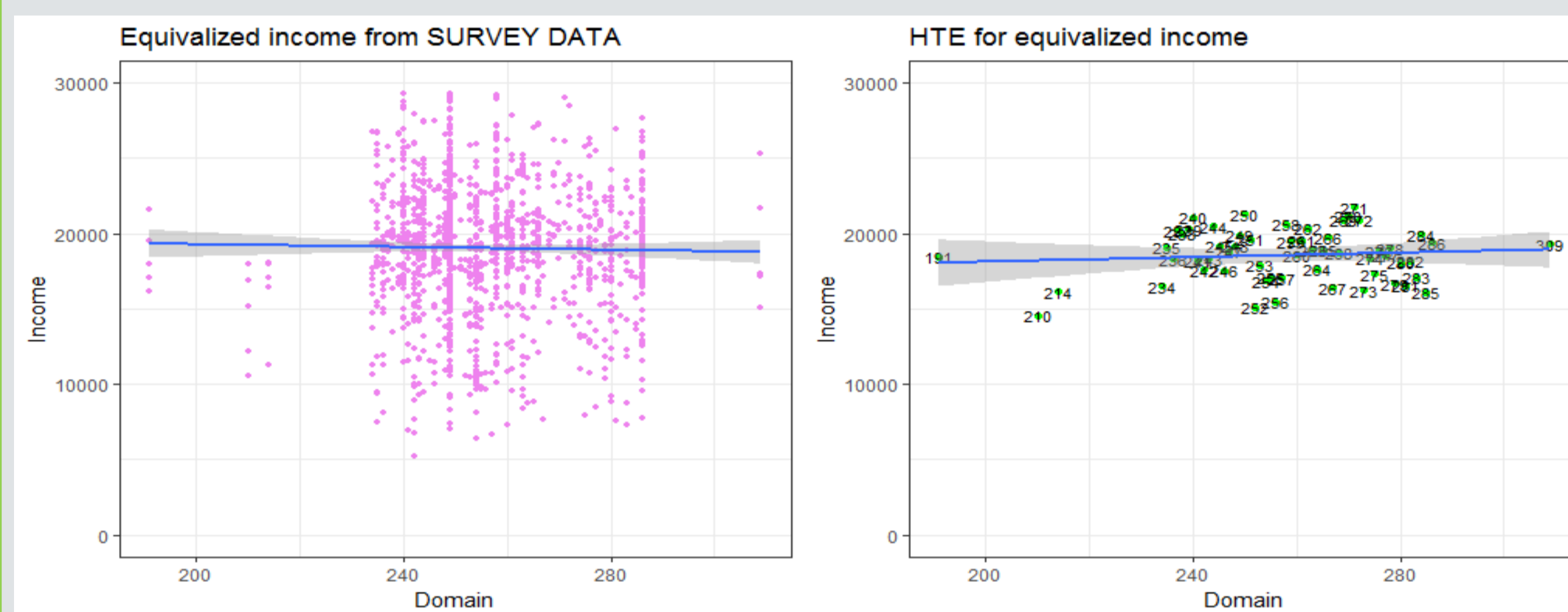
**Target territory:** Tuscany region, Italy.  
**Census.Rdata:** Data corresponding to 1,388,252 households.  
**Data\_Area\_Level.Rdata:** Contains area level information that can be used as covariate info at SAE model.  
**SurveyC.Rdata:** samples generated from the real EU-SILC 2008.

## USED ESTIMATORS:

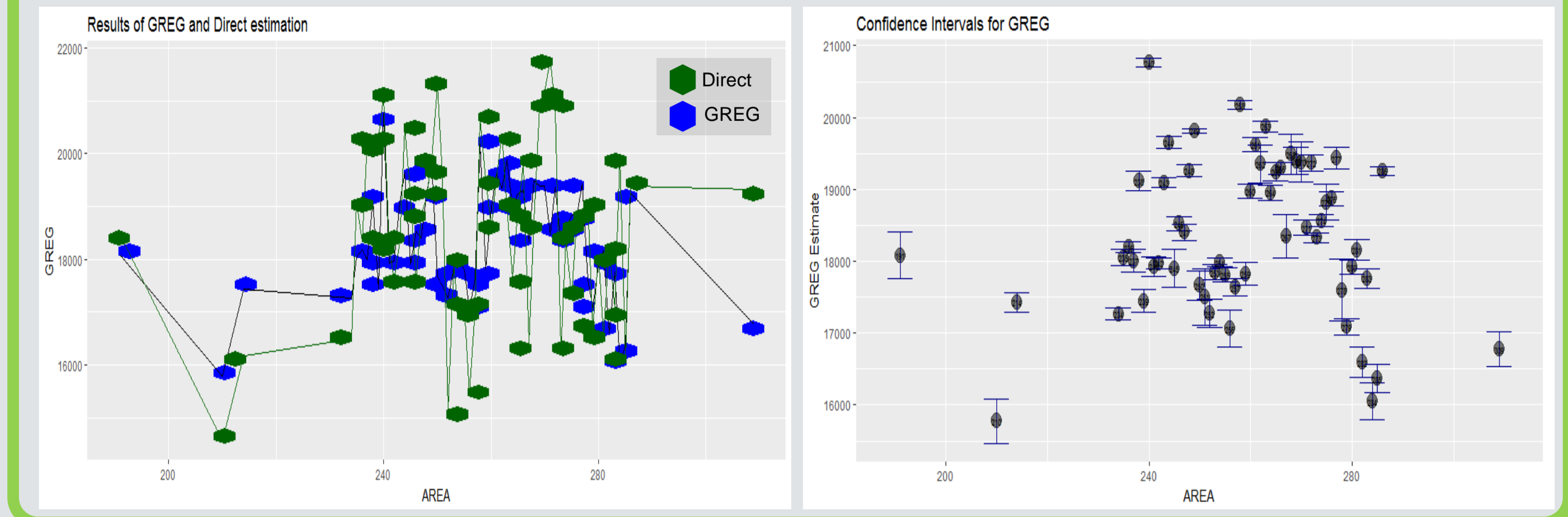
- Direct Estimator - Horvitz and Thompson Estimator (HT)
- SAE:
  - Generalized Regression(GREG)
  - Empirical Best Linear Unbiased Prediction(EBLUP)
  - Fay Herriot (FH-EBLUP)

## Horvitz and Thompson Estimator (HT)

- Is direct estimator.
- Is a general estimator for a population total
- Is an unbiased estimator under unequal sampling
- Can be used for any probability sampling plan.
- Both sampling with and without replacement.

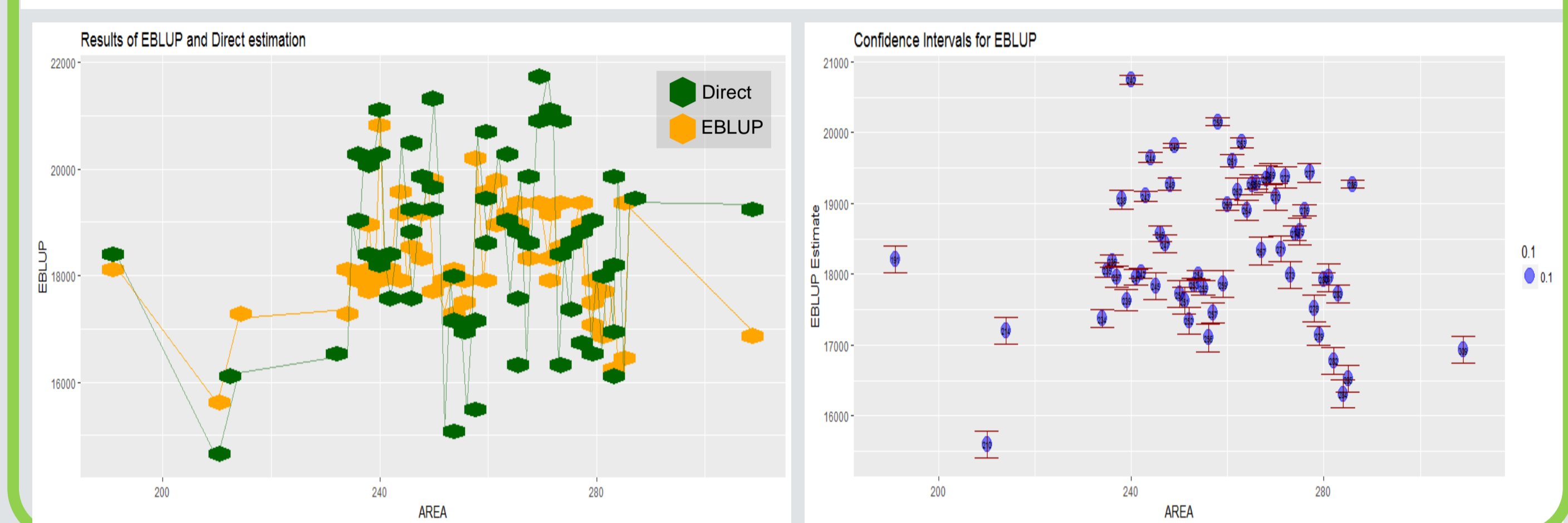


## Results of GREG Estimation

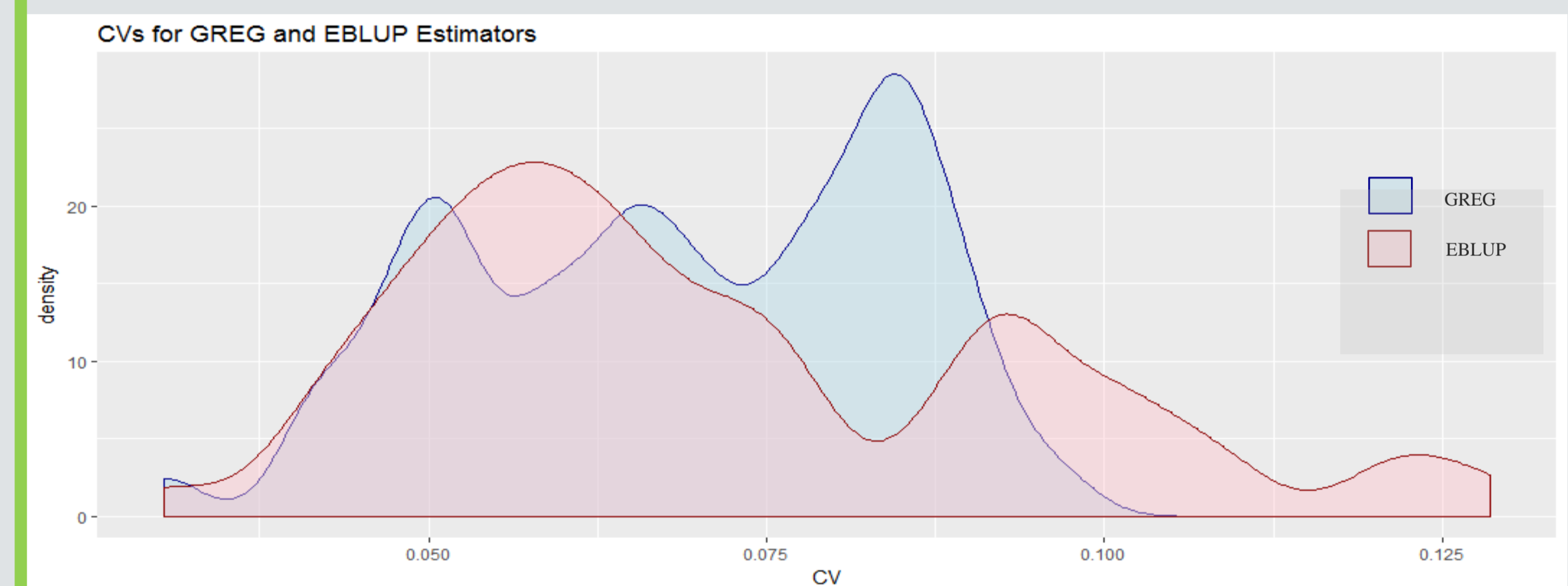


## Model based estimation - EBLUP estimator

- An efficient estimator under the *normality assumption* of the linear mixed model.
- Allows for random area effect and individual random effect.
- Is specified under the assumption of linearity of the relationship between the study variable and the auxiliary variable.
- Not design consistent and not robust to outliers.
- Out of sample predictions come only from area specific auxiliary variables.



## COMPARISON



- Generally there is not one the best estimator, it depends from case to case.
- In wordly, if we have enough sample size, direct estimator will be precise.
- In conclusion we can say that it's not happen in here.
- If we look the total graph, we can say EBLUP estimator is better totally because it has small CV.
- But by area we can choose GREG estimator or EBLUP estimator for the best estimator, it depends on the LLS.
- While GREG is *unbiased*, EBLUP is *biased*(doesn't approach to 0 with increasing sample size).

## MODEL ASSISTED ESTIMATION GENERALIZED REGRESSION(GREG)

We use indirect estimation, when direct estimation is not possible.

First we need the find significant covariates for using model of greg and eblup estimators.

P value shows us; "how probable the results were due to luck ". We found p-value 2e-16 and it means there is %2e-16 percent of probability that the results were due to random chance.

Because of that SM\_House, Work\_Status, Gender, Year\_Education, Single are significant and we will use this covariates to finding greg and eblup estimators.

- Have only 1 regression model with fixed effect(results of small-area estimators can overlook the area effects).
- Is asymptotically design-unbiased and consistent, but can be sensible to extreme values of inclusion probabilities.
- Allows for survey weighting of outlying observations, but this does not guarantee protection against outlying observations.

## Fay – Herriot Area Level Model

