Living Conditions and Poverty in Italy - 2017

SMALL AREA ESTIMATIONS AND MACHINE LEARNING METHODS

SAMPIEU Final Event | VIVIN VINOD | v.vinod@studenti.unipi.it
19/07/2021

INTRODUCTION

POVERTY – A BROADER PICTURE

Analysing Poverty

- Multidimensional nature
- Poverty indicators
- What we will do
  - Indicator Analysis
  - SA Methods
    - HCR
    - Mean Income
  - Machine Learning models

INDICATORS AND ESTIMATION METHODS

DIRECT

SA METHODS

The Head Count Ratio

- Measures the proportion of the population that is counted as poor
- Given by

$$P_0 = \frac{N_P}{N}$$

- By poverty line (z) and the actual income (y)

$$P_0 = \frac{1}{N} \sum_{i=1}^{N} I(y_i < z)$$

ARP Gap

- Measures the number of individuals on average (or median) who are below the poverty line and reports this as a percentage of the poverty line.
- It is the difference between the poverty line (z) and the actual income (y)

$$G_i = (z - y_i) I(y_i < z)$$

$$P_1 = \frac{1}{N} \sum_{i=1}^{N} \frac{G_i}{z}$$

The Quintile Share Ratio

- S80/S20 ratio

$$QSR = \frac{\sum_{j \in \text{S80}} w_j \cdot Y_j}{\sum_{j \in \text{S20}} w_j \cdot Y_j}$$

The Gini Index

- Demonstrates a degree of inequality in a distribution of income/wealth

$$G = \frac{1}{n} \cdot (n + 1 - 2\frac{\sum_{i=1}^{n}(n + 1 - i)Y_i}{\sum_{i=1}^{n} Y_i}$$

$$\text{GINI Index} = G \times 100$$

The Horvitz Thompson Estimator

- Design-consistent and unbiased direct estimator
- Reliability of the HT estimator
  - Planned domain
  - Disaggregation level
  - Large sample size
- HT direct estimator for household income:

$$\bar{y}_i^{HT} = \frac{1}{n} \sum_{k=1}^{n} a_k y_k$$

- Degree of precision – CV $= \frac{\sigma}{\mu}$

The Fay-Heriot Model

- Auxiliary data and Model based approach
- Summary data on area level vis-à-vis estimates
- FH for household income

$$\hat{y}_i^{dir} = x_i^T \beta + u_i + e_i$$

- Best linear unbiased predictor

$$\tilde{y}_i^{EBLUP} = \gamma_i \hat{y}_i^{dir} + (1 - \gamma_i) x_i^T \tilde{\beta}$$

The data

A BRIEF OUTLINE OF THE DATASETS

From sourcing data to source code

- Cross sectional NUT2 level EUSILC-2017 data on Italy
- Administrative data from ARCHIMED from ISTAT
  - Municipality level
  - Selected regions
  - Equivalised household income quantile values
  - Type of employment
  - Age range
- Source code in R using laeken, sae and python using minisom

KEY VARIABLES:

- *Equivalised household income*
- *Provinces*
- *Sampling weights*

Empirical Analysis and Results

AND GRAPHS

Indicators and Direct Estimates

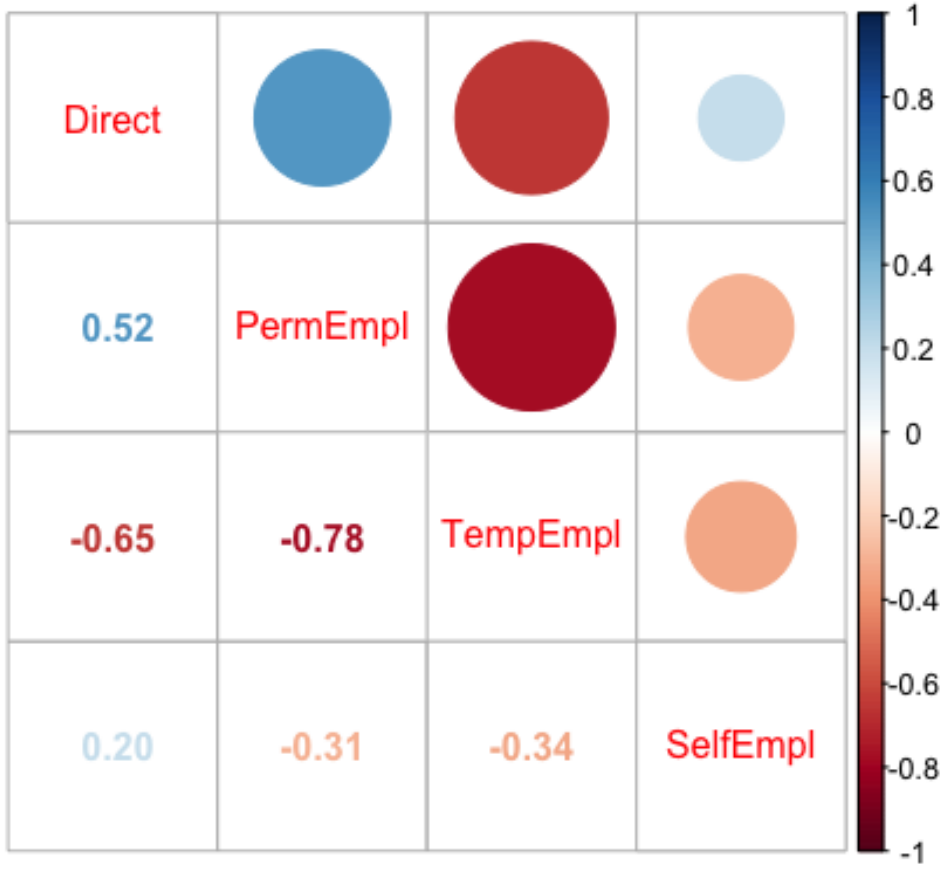
- Overall Estimates
  - Poverty line: €9997.2
  - HCR: 20.7
  - Gini: 0.34
  - QSR: 6.3
  - Poverty gap: 29.3
- HCR by Sex
  - Males: 18.2
  - Females: 25.3

Age	HCR
[0,20]	37.664
(20,40]	30.405
(40,60]	28.224
(60,80]	25.837

- Provincial Estimates
  - HCR: 3.2% - 49.32% (Median–15.87%)
  - Mean income: € 10854- € 28392 (Median–€ 18823)
- CVs
  - 16.6% - 33.3% (unplanned domains)
  - >33% outlier
  - Thus, precision of estimates not unconditional
- Inadequacy

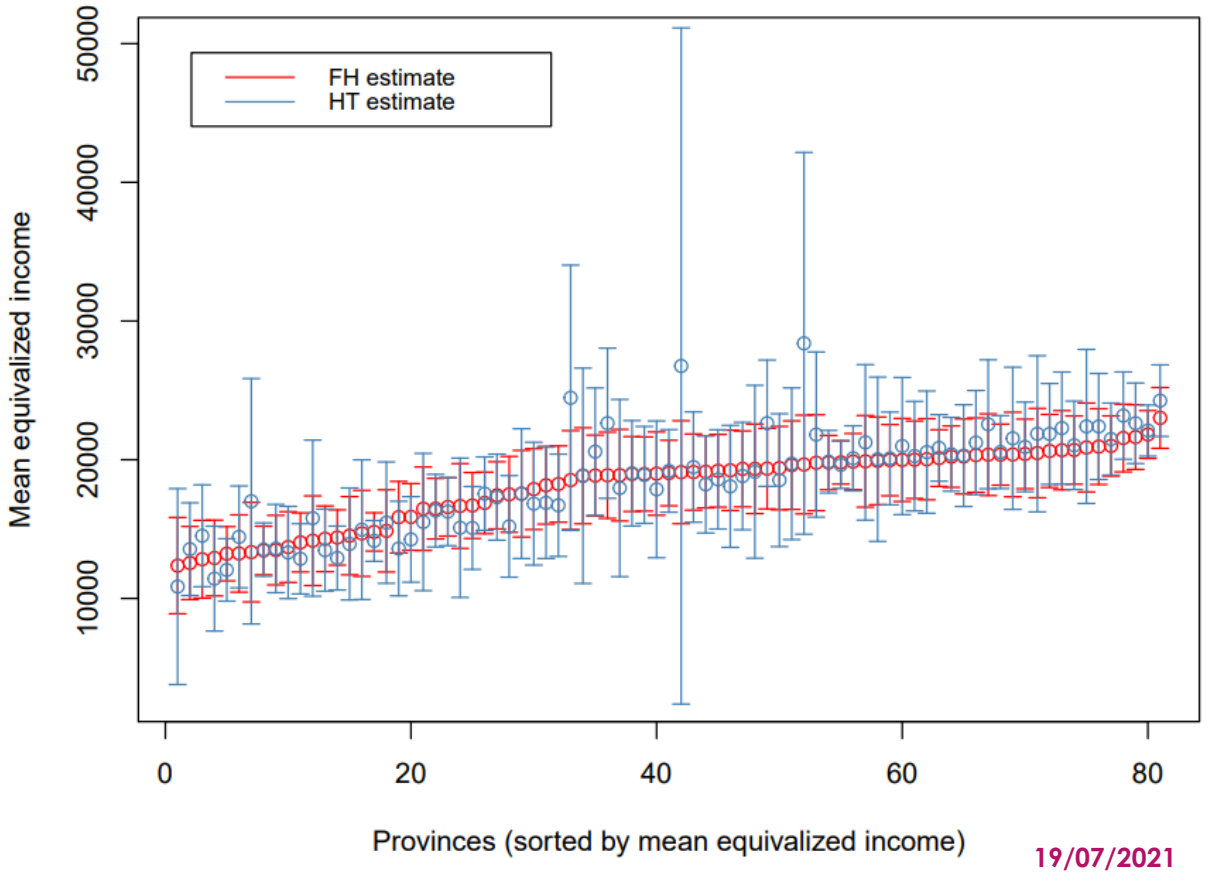
Model Based Estimates

- Auxiliary variables:
  - Employment type
  - Adult (35-64)
  - Active population (15-64)
  - Household composition

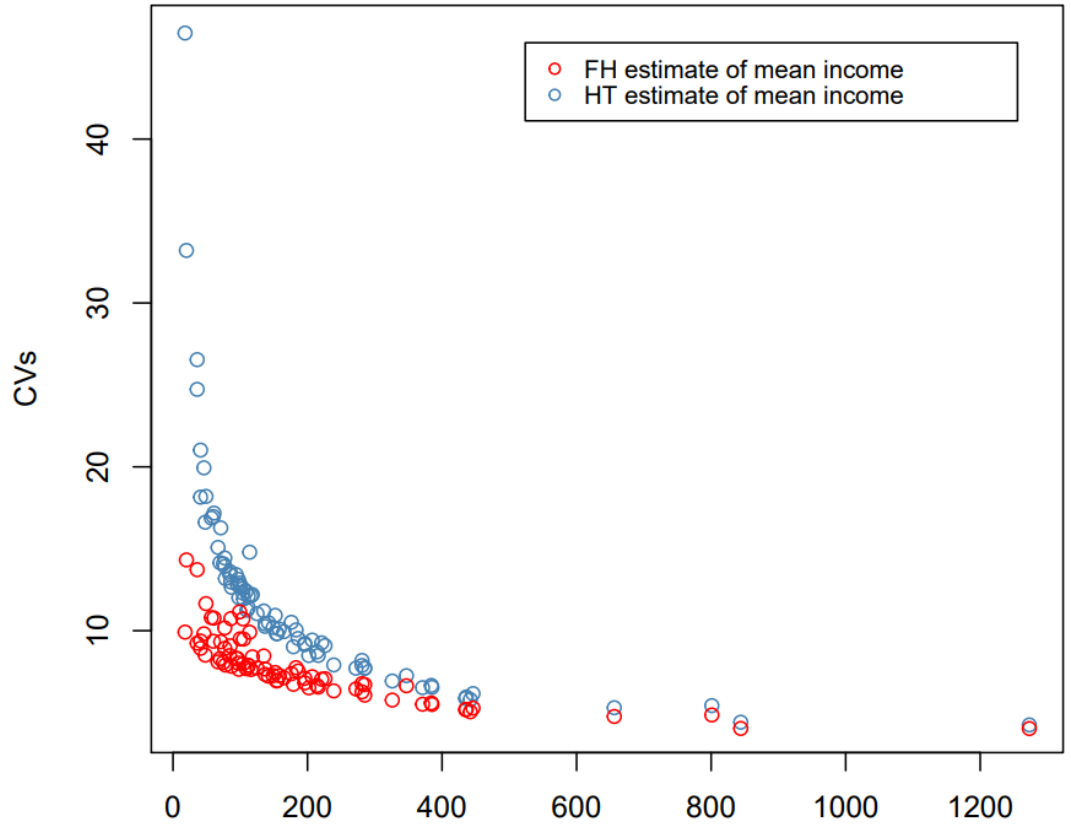


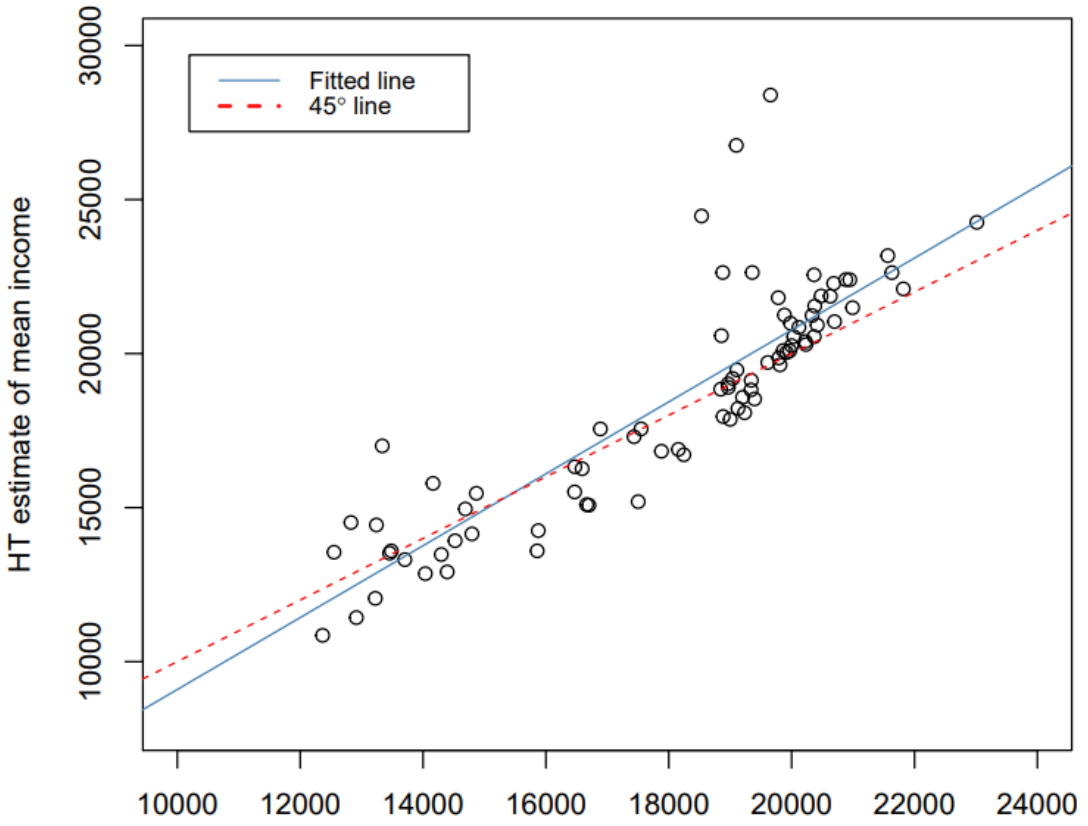
Mean Income – Model Estimates

- Two models
  - different types of employment + share of adults
  - active population + single households + types of employment
- Direct: €10854 - €28392 (€18823)
- FH: €12365 - €23014 (€19040)

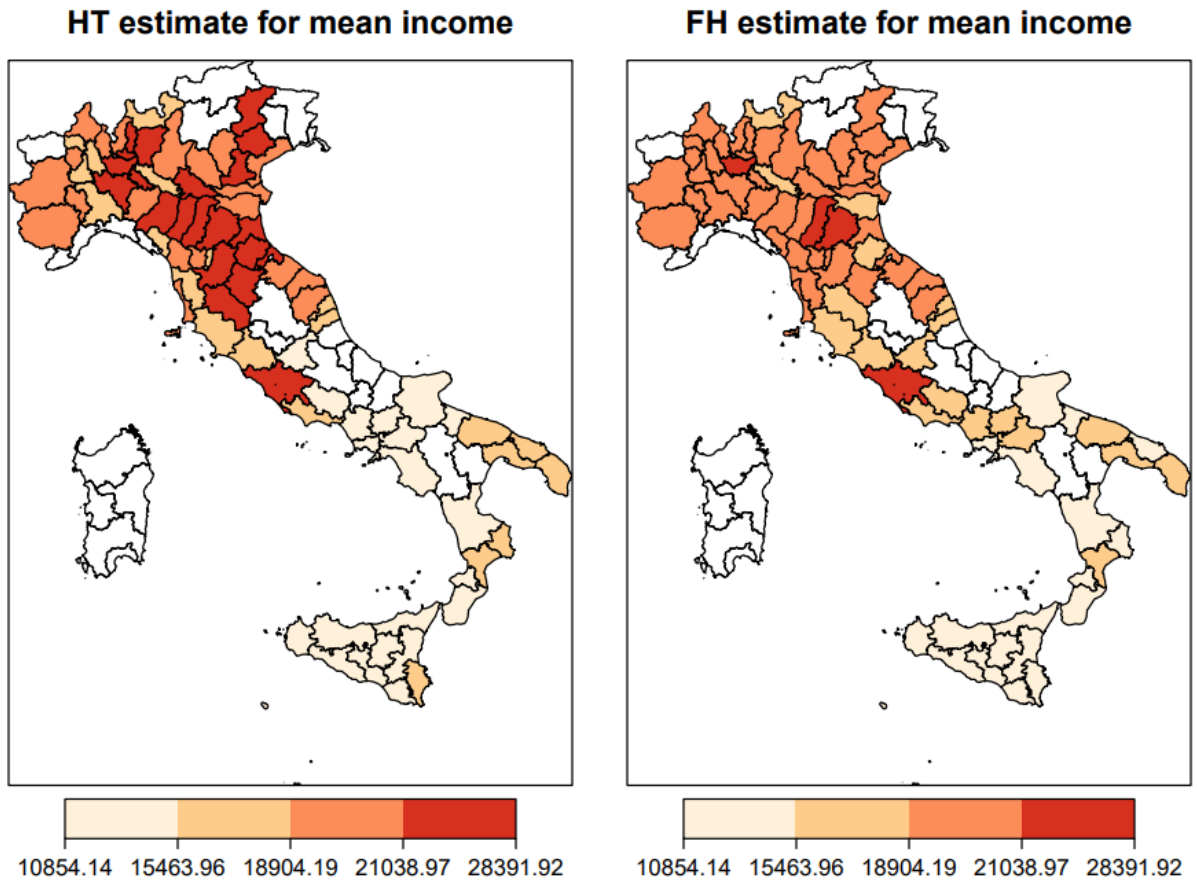


Mean Income – CV and bias



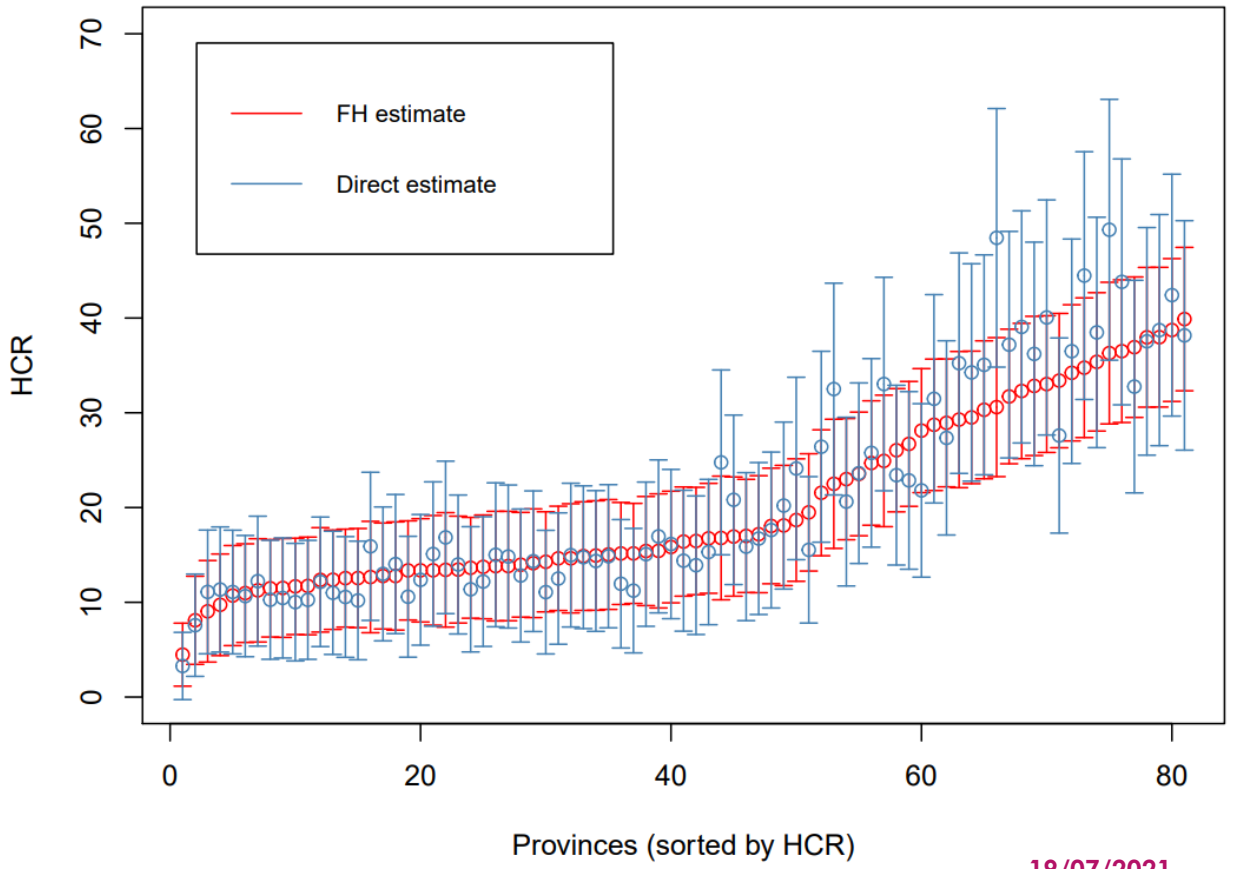


Mean Income - Maps

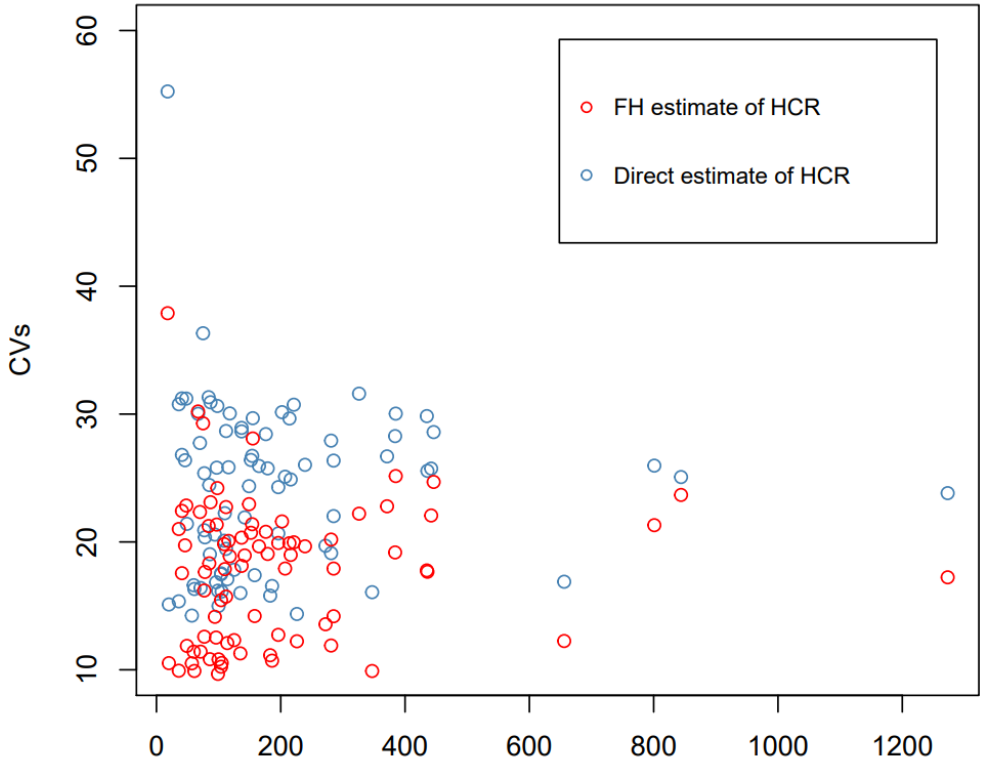


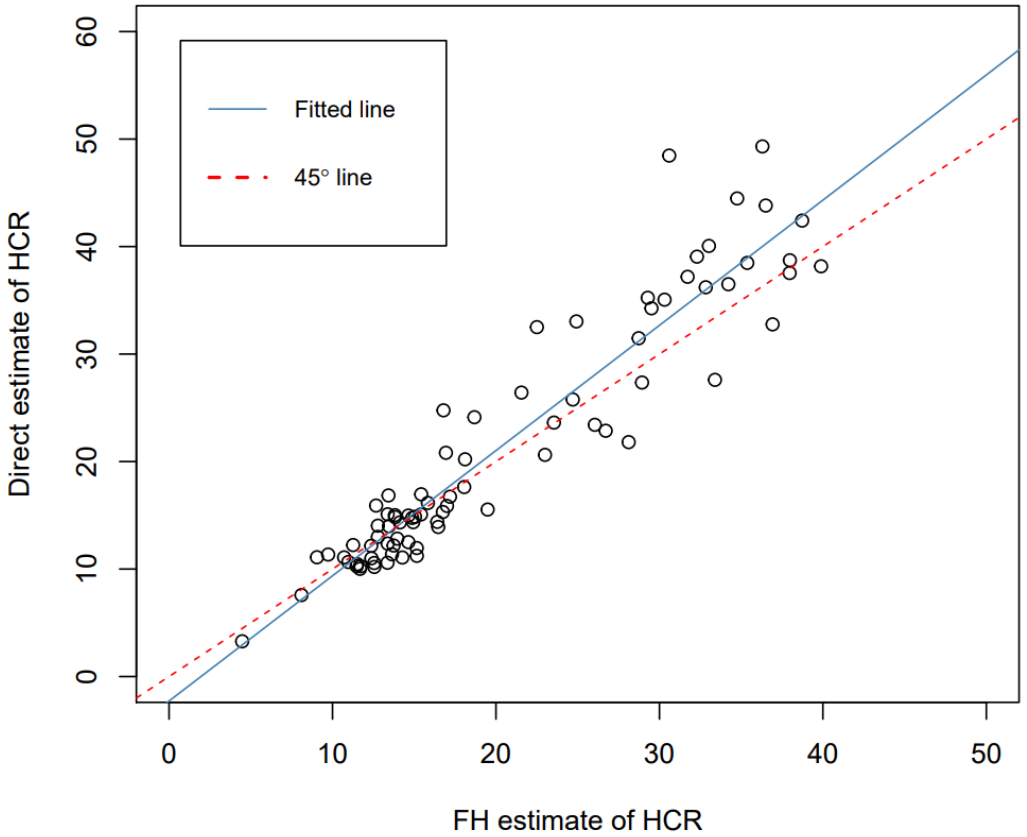
HCR

- Direct: 3.2 - 49.32 (15.87)
- FH: 4.48 - 39.89 (16.41)

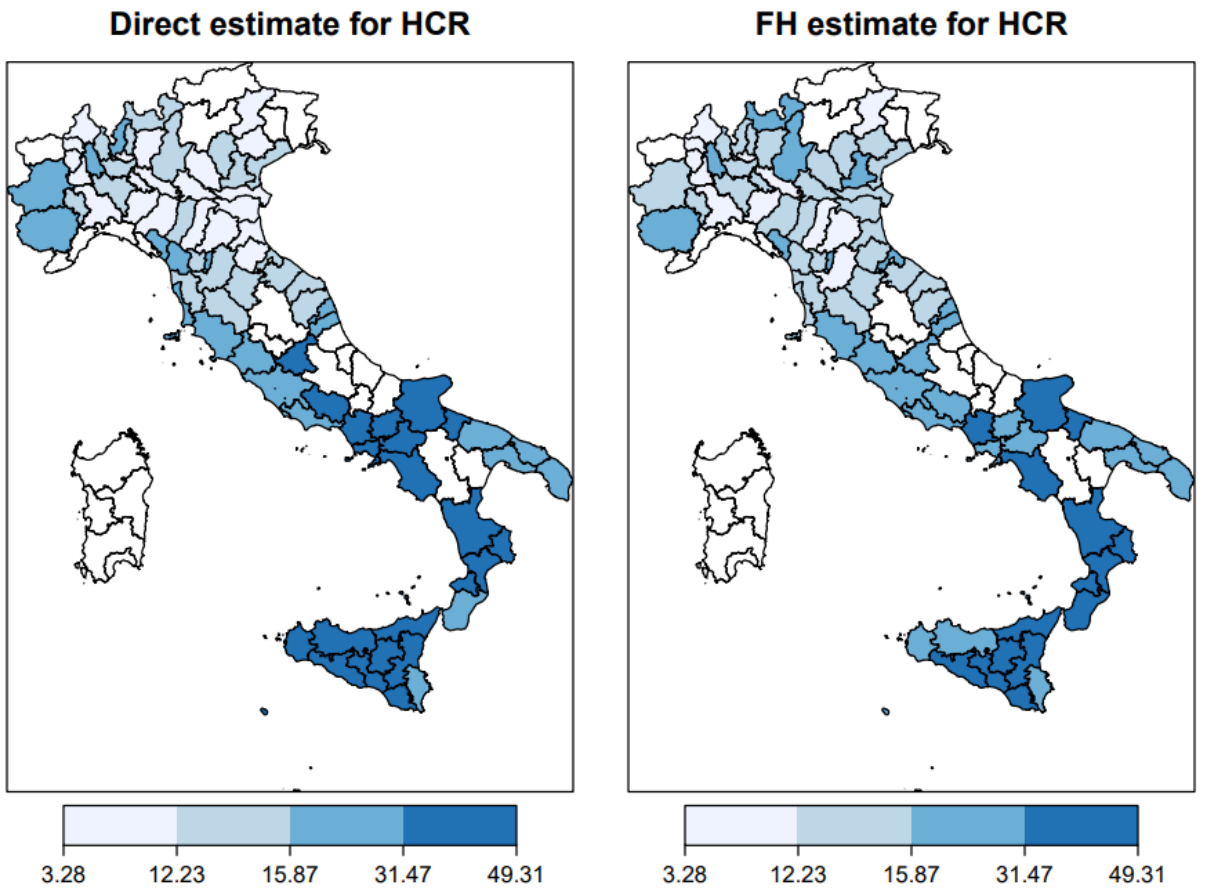


HCR – CV and bias





HCR - Maps

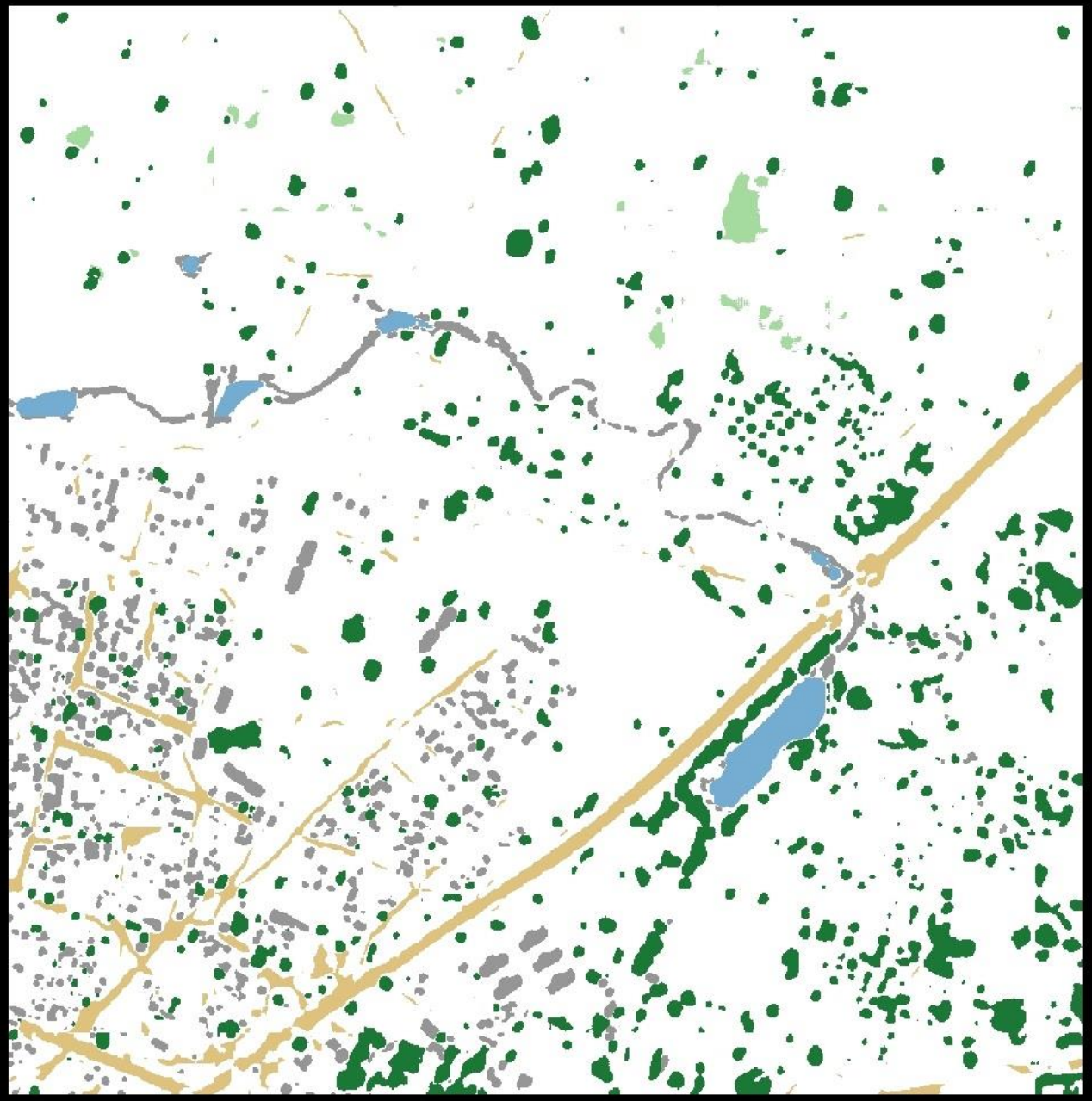

Direct estimate for HCR
3.28
12.23
15.87
31.47
49.31
FH estimate for HCR
3.28
12.23
15.87
31.47
49.31

Why Machine Learning

- Augmented method
- Requires very little prior experience
- Scalability
- Versatility
- No issues of homoskedasticity or heteroskedasticity etc
- No bounds on data size, type, distribution
- Ground-up construction of indicators
- Wide range of applications

Satellite Image Segmentation

- Deep learning U-NET model
- Supervised categorization, labelled dataset
  - Roads
  - Buildings
  - Crops
  - Trees
  - Water bodies

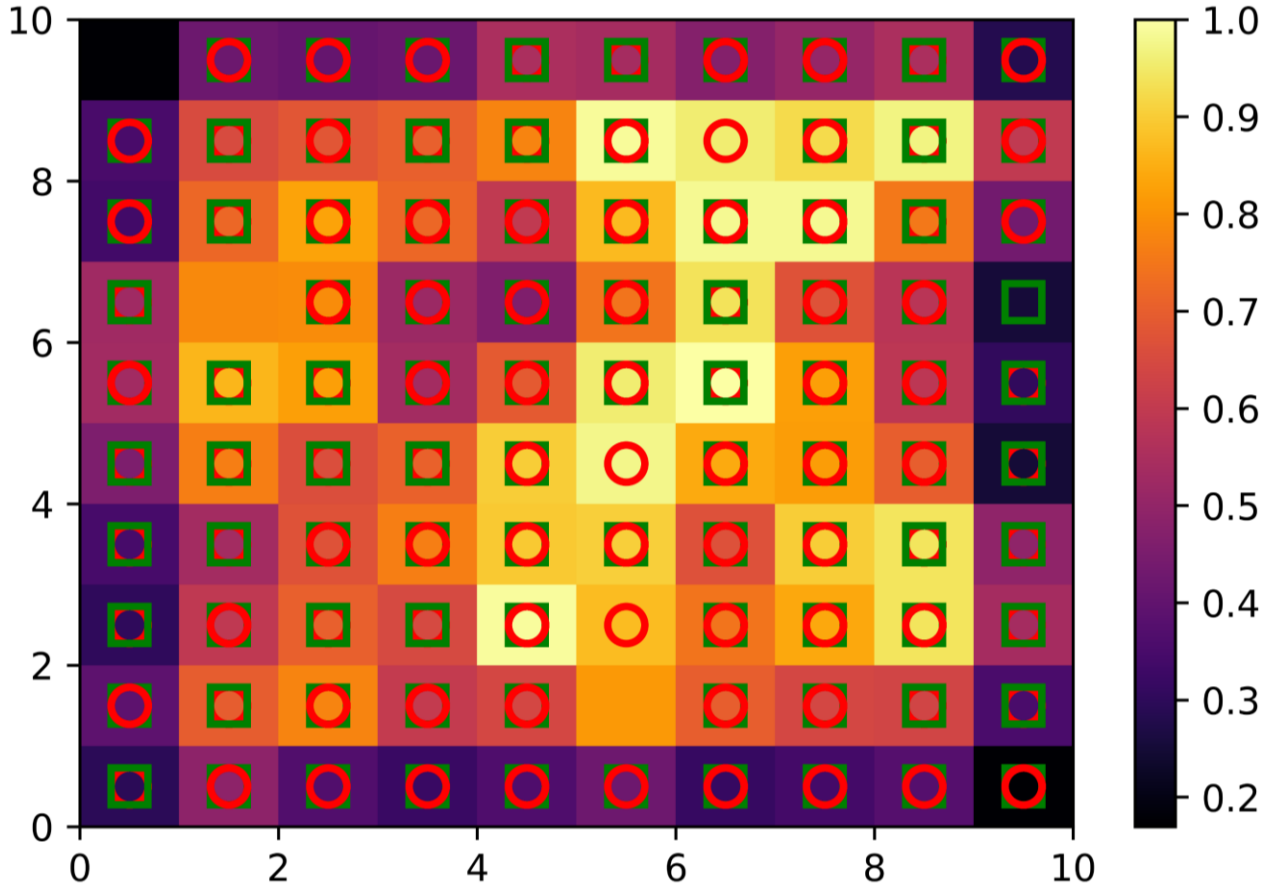


Satellite Image Segmentation – Why?

- Assessing land cover
- Measure of urbanization
- Creating metrics
- Augmented Indicators as a new set of complex indicators
- Challenge – quantifying qualitative results

Clustering - SOMs

- Clustering
- Unsupervised, unlabelled
- Output analysis of outliers
- Interpretations
- Mapping to quantifiable results
- Specificity of policy
- Degree of similarity



Limitations

AND IMPROVEMENTS

Sample Data Errors

- Coverage errors
- Non-response errors
- Measurement errors
- Sampling and non-sampling error

HT vs FH Measures

HT

- Sample data
- Sample size reliability / level of dis
- Limited precision for small sample size

FH

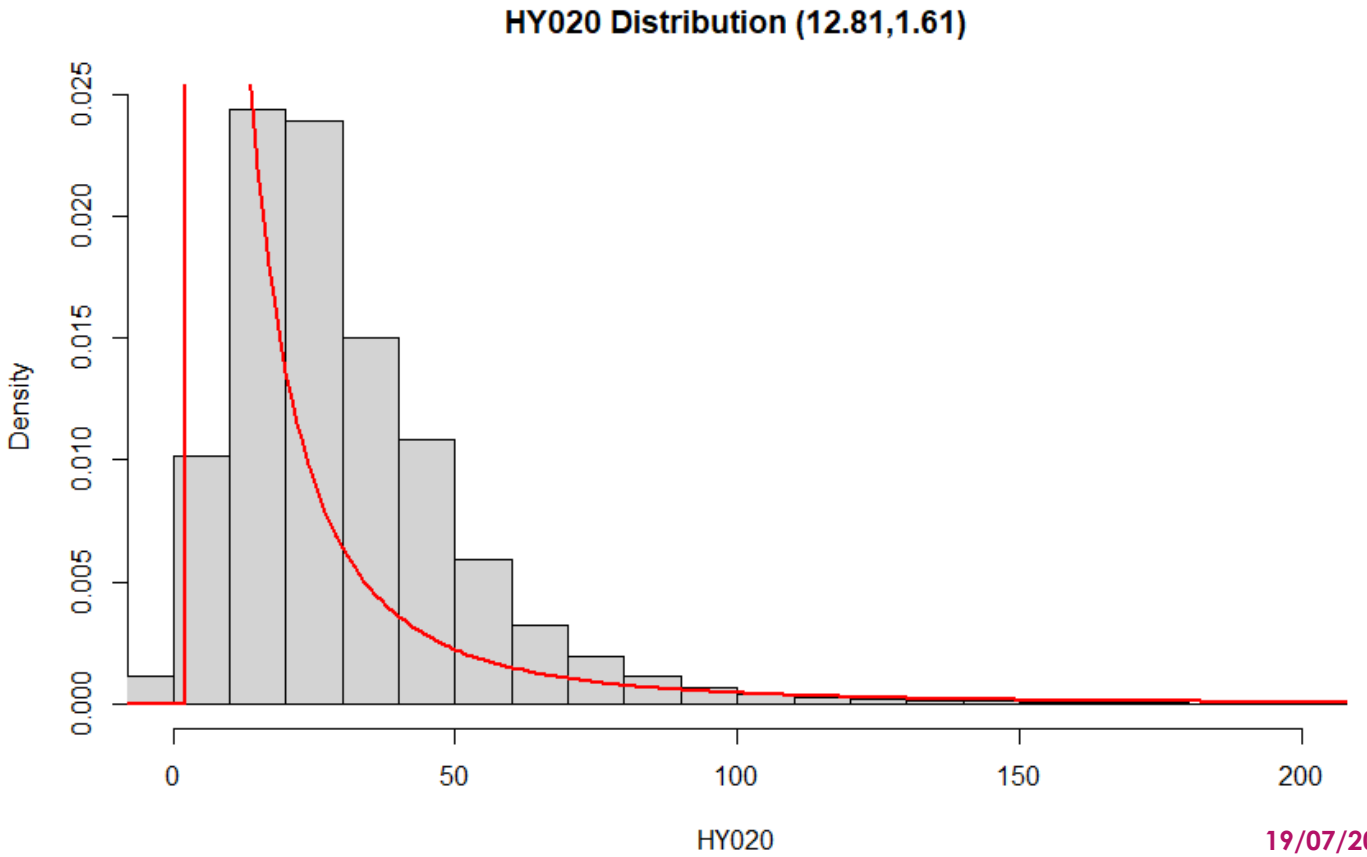
- Sample data + auxiliary data
- Enhancement – predictive power of auxiliary variables
- Overcome small sample size problems

Policy Recommendations

TAKE HOME POINTERS

Government Recommendations

- Stage I – Planning
- Stage II – Reporting
- Stage III – Recommending and Reporting



Welfare Recommendations

- Transparency
- Ease of application
- Hard-to-reach population access

Intervention Recommendations

- Lowering inflation
- Public sector investment
- Housing supply
- Free trade agreements

THANK YOU

REFERENCES

- Agustin, Diana et al. "Analysis of Multidimensional Poverty Indicators in Indonesia with Association Rules". In: Jurnal Ekonomi Pembangunan 18 (Dec. 2020), page 136. doi: 10.22219/jep. v18i2.14244
- Foster, James. "Poverty Indices". In: June 2006, pages 41–65. isbn: 978-1-4020-7850-7. doi: 10. 1007/0-387-29748-0_4
- Jarman, B, Townsend, P, and Carstairs, V. "Deprivation indices". In: BMJ (Clinical research ed.) 303 (Sept. 1991), page 523. doi: 10.1136/bmj.303.6801.523-a
- Jenkins, Stephen and Cappellari, Lorenzo. "Summarizing multiple deprivation indicators". In: (Sept. 2007)
- Kovacevic, Milorad and Calderon, M. "UNDP's Multidimensional Poverty Index: 2014 Specifications". In: (Oct. 2016). doi: 10.13140/RG.2.2.11827.89127
- NOVAK, I.M. "Deprivation Energy Poverty Indicators". In: Demography and social economy (Oct. 2020), pages 158–172. doi: 10.15407/dse2020.03.158
- Radford, J and Campbell, D. "Deprivation indices". In: BMJ (Clinical research ed.) 303 (Nov. 1991), page 857. doi: 10.1136/bmj.303.6806.857