EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS

2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS
2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union

Indirect estimation of selected characteristics of the working and unemployed population in the functional areas of provincial capitals

Tomasz Józefowski[1,2], Kamil Wilak [1,2]

[1] Statistical Office in Poznań
[2] Poznań University of Economics and Business, Poland

Outline of the presentation

Motivation

Methodology

Empirical study

Conclusions

Literature

Motivation

- Growing information needs of data users.
- Limitations of the direct estimator.
- Costs of sample surveys.
- The need to reduce the respondent burden.
- Established literature on indirect estimation.
- Practical applications of SPREE.

Methodology

- Structure Preserving Estimators (SPREE) are a generalised class of synthetic estimators, i.e. those that rely on information from direct estimates.
- SPREE estimators can be used to estimate totals for small areas during intercensal periods or for more detailed domains.
- SPREE estimators rely on adjusted counts withing the contingency table, which are obtained by applying the method of Iterative Proportional Fitting (IPF).

Methodology

	Domain 1	Domain 2	...	Domain K	
Area 1	X_{11}	X_{12}	...	X_{1K}	$Y_{1.}$
Area 2	X_{21}	X_{22}	...	X_{2K}	$Y_{2.}$
⋮	⋮	⋮	⋱	⋮	⋮
Area A	X_{A1}	X_{A2}	...	X_{AK}	$Y_{A.}$
	$Y_{.1}$	$Y_{.2}$	...	$Y_{.K}$	$Y_{..}$

- SPREE consists in updating counts inside the contingency table so that they add up to known marginal totals.
- Input counts inside the contingency table can come from a census or an administrative register, while reliable direct estimates from a survey (e.g. LFS) are used as marginal totals.

Methodology

- Y_{ak}, X_{ak} – denote values of the response variable and the proxy variable in a cell of a two-way contingency table, where a denotes the identifier of a small area, while j denotes levels of the grouping variable.
- Let $logY_{ak}$, $logX_{ak}$ be expressed by a log-linear model:

$$logY_{ak} = \alpha_0^Y + \alpha_a^Y + \alpha_k^Y + \alpha_{ak}^Y, \tag{1}$$

$$logX_{ak} = \alpha_0^X + \alpha_a^X + \alpha_k^X + \alpha_{ak}^X, \tag{2}$$

where:

- $\alpha_0^Y = \frac{1}{AK} \sum_a \sum_k \log Y_{ak}$ – general effect,
- $\alpha_a^Y = \frac{1}{K} \sum_k \log Y_{ak} - \alpha_0^Y$ – effect of area a,
- $\alpha_k^Y = \frac{1}{A} \sum_a \log Y_{ak} - \alpha_0^Y$ – effect of k-th level of the grouping variable,
- $\alpha_{ak}^Y = \log Y_{ak} - \alpha_a^Y - \alpha_k^Y - \alpha_0^Y$ – interaction effect,

for $a = 1, \ldots, A$ and $j = 1, \ldots, J$.

Methodology

- SPREE estimator is based on the assumption that:

$$\alpha_{aj}^{Y} = \alpha_{ak}^{X}. \tag{3}$$

- Assumption (3) can be relaxed by using a GLSM model and adopting the assumption about the existence of a proportional relationship between association structures of the response variable and the proxy variable. In this way we obtain a generalised SPREE (GSPREE):

$$\alpha_{ak}^{Y} = \beta \alpha_{ak}^{X}. \tag{4}$$

- The literature provides other modifications of this estimator as well as their applications.

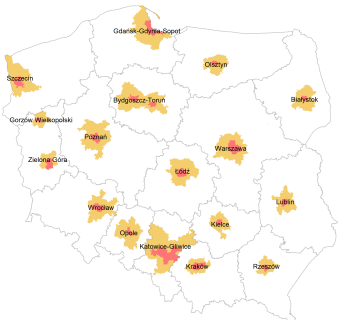
Empirical study

- **Data:**
  - aggregated statistical data based on administrative registers (2017),
  - Labour Force Survey (2017);
- **Domain:**
  - functional urban areas of provincial capital cities broken down by: (1) sex, (2) age groups (mobility age, non-mobility age);
- **Estimated parameter:**
  - the number of people in employment, unemployed, economically inactive,
  - labour market indicators (Economic Activity Rate, Employment Rate, Unemployment Rate);
- **Association structure:**
  - counts obtained from aggregated statistical data based on administrative registers,
- **Allocation structure:**
  - LFS-based direct estimates of marginal totals in the target contingency table;
- **Methods:**
  - CAL, $SPREE_{2,L}$, $SPREE_{3,L}$, $GLSM_{2,L}$,$GLSM_{3S,L}$.

The conference is partly financed by the European Union

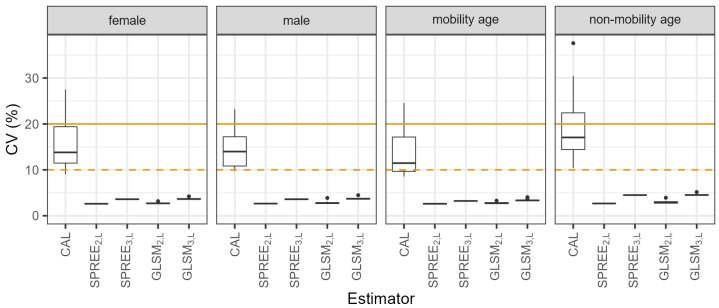
Empirical study

Functional urban areas of provincial capital cities selected for Integrated Territorial Investments (ITI) in the period 2014-2020



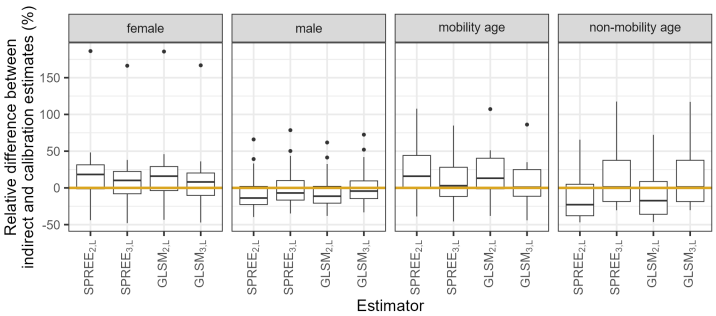
Empirical study

CV coefficients of unemployment rate estimates for functional urban areas of provincial capital cities



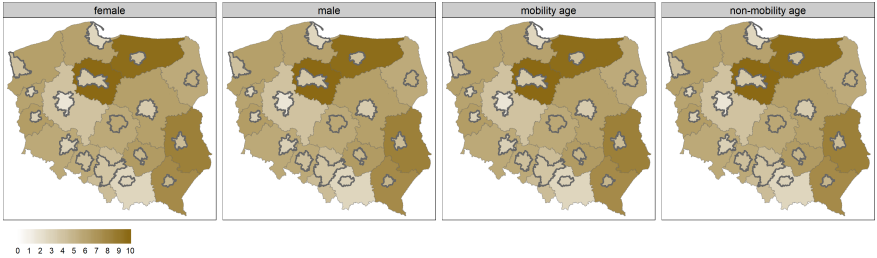
Empirical study

Relative differences between indirect and calibration estimates of the unemployment rate in functional urban areas of provincial capital cities



Empirical study

SPREE estimates of the unemployment rate in functional urban areas of provincial capital cities and the remaining parts of the provinces



The conference is partly financed by the European Union

Conclusions

- Estimates obtained by applying SPREE are consistent with direct estimates at higher levels of spatial aggregation (benchmarking).
- Compared with the direct estimator, SPREE estimators are characterised by better precision.
- Statistical data based on administrative registers seem to be a good source of auxiliary variables.
- The SPREE approach can be used to produce estimates for non-standard territorial domains, such as functional urban areas of provincial capital cities.

Literature

Józefowski, T. (2022). *Rynek pracy na obszarach funkcjonalnych miast wojewódzkich w świetle estymacji typu SPREE*. PhD thesis, Poznan University of Economics and Business.

Józefowski, T. Szymkowiak, M. (2013). *Zastosowanie estymatora typu SPREE w szacowaniu liczby osób bezrobotnych w przekroju podregionów*. Studia Oeconomica Posnaniensia, 1(10):120–136.

Luna Hernandez, A. (2016). *Multivariate Structure Preserving Estimation for Population Compositions*. PhD thesis, University of Southampton.

Purcell, N. J. Kish, L. (1980). Postcensal Estimates for Local Areas (Or Domains). *International Statistical Review / Revue Internationale de Statistique*, 48(1):3–18.

Rao, J. M. K. Molina, I. (2015). *Small Area Estimation*. John Wiley and Sons, New Jersey, 2nd edition.

Zhang, L. C. Chambers, R. L. (2004). Small area estimates for cross-classifications. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 66(2):479–496.

The conference is partly financed by the European Union

Thank you for your attention

EUROPEAN CONFERENCE ON QUALITY IN OFFICIAL STATISTICS

2024 ESTORIL - PORTUGAL

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

eurostat

The conference is partly financed by the European Union