*Indirect estimation
of labour market characteristics
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

Poznań, 18-19.11.2024

Plan prezentacji

Introduction

Methodology

Empirical study

Conclusions

Literature

Introduction

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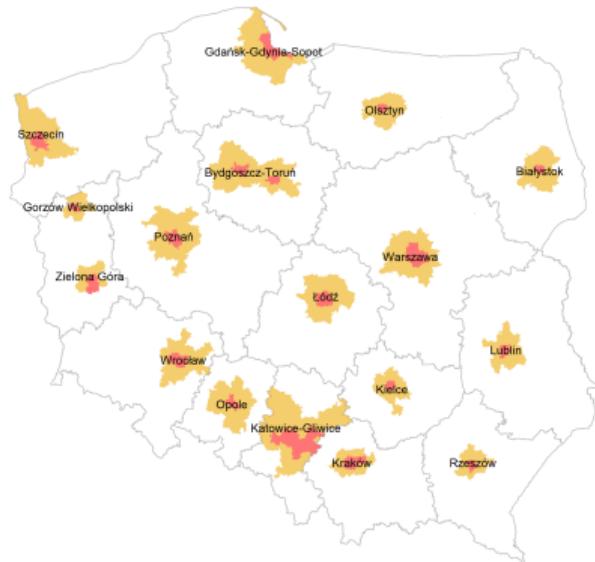
Introduction

Aims

- ▶ Estimation of labour market indicators in functional urban areas of provincial capital cities.
- ▶ Use of SPREE-type estimators to improve estimation quality.

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.

Statistical Office
in Poznań

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.

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}	$Y_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
Area A	X_{A1}	X_{A2}	...	X_{AK}	$Y_{A.}$
	$Y_{.1}$	$Y_{.2}$	...	$Y_{.K}$	

▶ SPREE estimators rely on adjusted counts within the contingency table, obtained by applying the Iterative Proportional Fitting (IPF) method, so that they add up to known marginal totals.

▶ Input counts $X_{a,k}$ 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 $Y_{a.}$ and $Y_{.k}$ ($a = 1, \ldots, A, k = 1, \ldots, K$).

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 $k = 1, \ldots, k$.

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 employed, unemployed and economically inactive persons,
  - 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, $SPREE_{2,L}$, $SPREE_{3,L}$, $GLSM_{2,L}$, $GLSM_{3S,L}$.

Empirical study

One of the association-allocation structures used in the empirical study.

	Functional urban area			Other parts of the province			
	employed	unemployed	inactive	employed	unemployed	inactive	
mobility-age							
dolnośląskie	$X_{1,1}$	$X_{1,2}$	$X_{1,3}$	$X_{1,4}$	$X_{1,5}$	$X_{1,6}$	$Y_{1,.}$
kujawsko-pomorskie	$X_{2,1}$	$X_{2,2}$	$X_{2,3}$	$X_{2,4}$	$X_{2,5}$	$X_{2,6}$	$Y_{2,.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
zachodniopomorskie	$X_{16,1}$	$X_{16,2}$	$X_{16,3}$	$X_{16,4}$	$X_{16,5}$	$X_{16,6}$	$Y_{16,.}$
non-mobility age							
dolnośląskie	$X_{17,1}$	$X_{17,2}$	$X_{17,3}$	$X_{17,4}$	$X_{17,5}$	$X_{17,6}$	$Y_{17,.}$
kujawsko-pomorskie	$X_{18,1}$	$X_{18,2}$	$X_{18,3}$	$X_{18,4}$	$X_{18,5}$	$X_{18,6}$	$Y_{18,.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
zachodniopomorskie	$X_{32,1}$	$X_{32,2}$	$X_{32,3}$	$X_{32,4}$	$X_{32,5}$	$X_{32,6}$	$Y_{32,.}$
	$Y_{.,1}$	$Y_{.,2}$	$Y_{.,3}$	$Y_{.,4}$	$Y_{.,5}$	$Y_{.,6}$	

Empirical study

SPREE estimates of the unemployment rate in functional urban areas of provincial capital cities and the remaining parts of the provinces, 2017
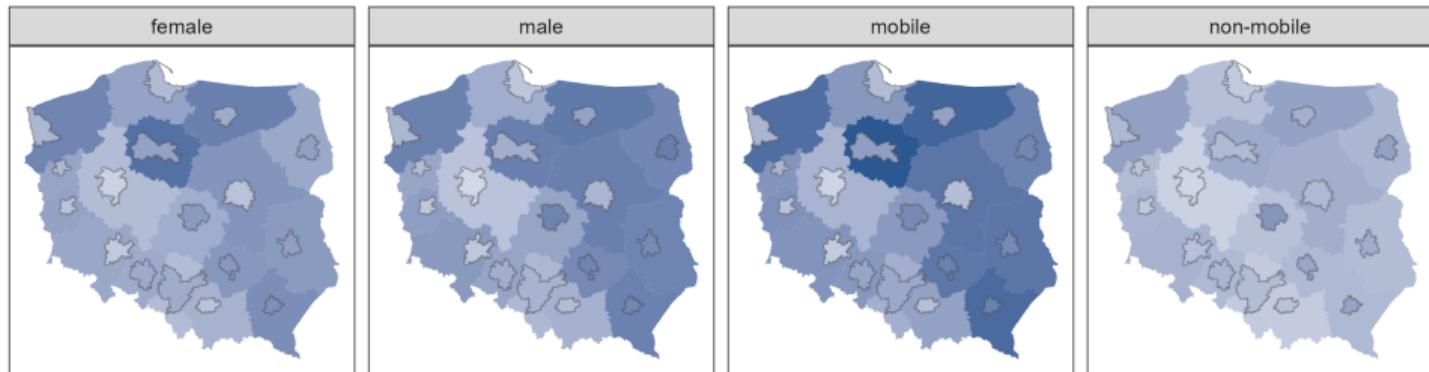
Conclusions

▶ Estimates obtained by applying SPREE are consistent with direct estimates at higher levels of spatial aggregation.

▶ 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.
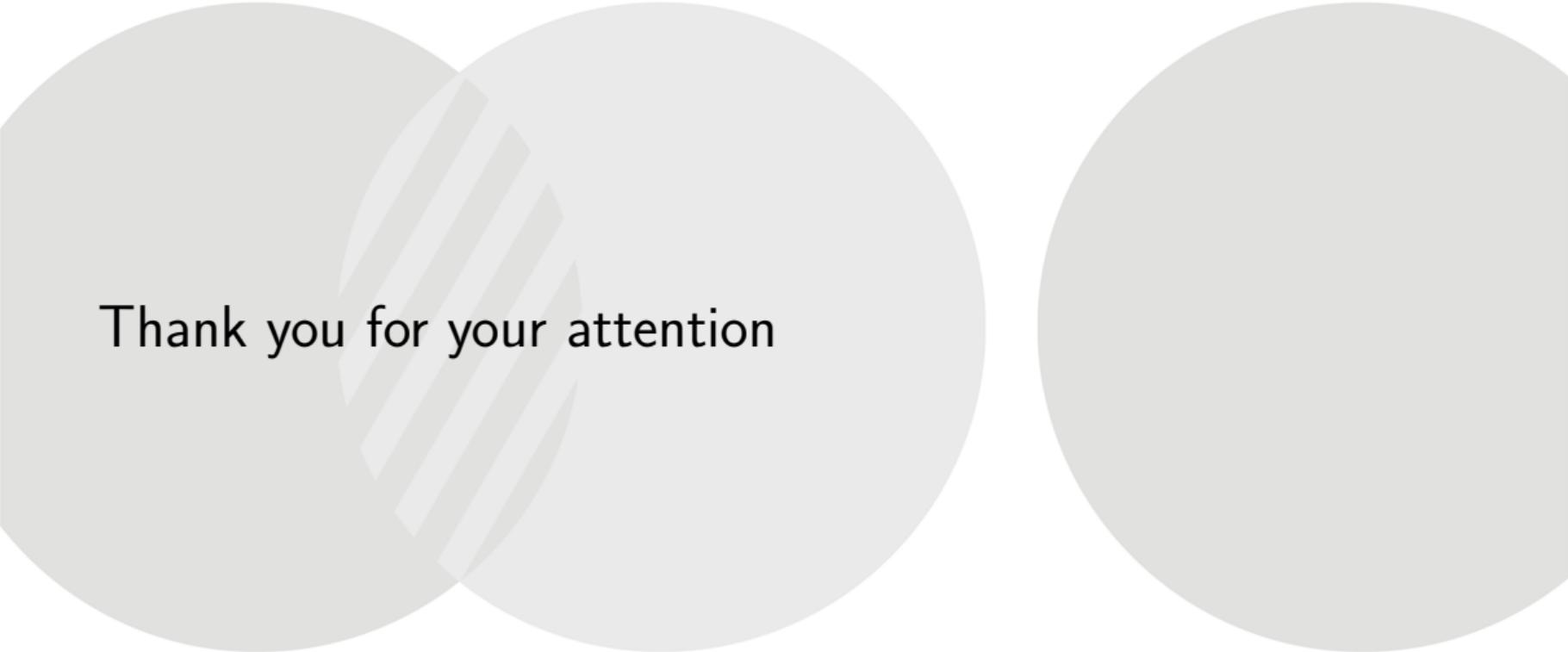
Thank you for your attention