

# Fay-Herriot model

Andrzej Surma

22 stycznia 2019

# Potrzebne dane z HHP

- kod GUS gminy gospodarstwa
- wielkość gospodarstwa domowego
- liczba dzieci (0-14), starszych (60+) w gospodarstwie
- dochód netto gospodarstwa
- klasa wielkości miejscowości
- wiek i płeć członków rodziny
- typ budynku
- modelowana wartość ważona kryteriami panelowymi

Do kalibracji wyników, niezależnie potrzebujemy informacji o raportowanych oficjalnie wartościach na poziomie województwa

## Selekcja zmiennych

- analiza korelacji
- analiza graficzna np. boxplot
- analiza regresji

## OLS

Zestaw zmiennych różnicujących średnie wydatki gospodarstwa, penetrację

## Dane geo o gminach

## Uwagi

- w geo: kody GUS ulegają aktualizacji każdego roku
- w panelu: brak aktualizacji, kody niesfunkcjonujące

## Macierz dystansów

poszukujemy podobnych gmin pod względem cech, które różnicują średnie wydatki gospodarstwa, penetrację

## Sąsiedztwo gmin i powiatów

geograficzne ujęcie bliskości

# Wiązkowanie gmin

- dla każdej gminy definiujemy dla niej wiązkę gmin podobnych, która może zrzeszać podobne pod względem zdefiniowanych cech gminy z sąsiednich powiatów. Nie jest wymagane, aby wiązka była 'łukowo spójna', tzn. aby można było połączyć gminy z wiązki linią
- grupa gmin podobnych - ważona metryka Gowera; wagi proporcjonalne do wyników metody selekcji zmiennych; Q3 dystansów jako cutpoint
- podobieństwo - macierz odległości oraz sąsiedztwo geograficzne
- uwierzytelnienie estymatora średniej wartości wydatków
- brak możliwości wnioskowania na poziomie gmin na podstawie HHP
- średnia oraz wariancja wartości wewnątrz wiązki

# Model Faya-Herriota

$$y_i = x_i^T \beta + v_i + e_i$$

$$\hat{\mu}_i = \hat{\gamma}_i \hat{y}_i + (1 - \hat{\gamma}_i) \bar{X}_i^T \hat{\beta}$$

- $i = 1, \dots, n$  - analizowane obszary - gminy
- $x$  - cechy wyjaśniające;  $y$  - zmienna wyjaśniana
- $v$  - element losowy na poziomie obszaru
- $\beta$  - estymowane współczynniki
- $\gamma$  - komponent wariancyjny z modelu FH

## sae::eblupFH

Modelowanie średniej wartości zakupów per gospodarstwo w gminie. Estymacja współczynników modelu odbywa się tylko na części gmin, dla których liczba raportujących gospodarstw w odpowiednich wiązkach jest  $\geq 30$ .



## Ostatnia faza

Za pomocą ważenia kwadratowego kalibracja wyników do wartości oficjalnie raportowanych przez HHP