

An index-based approach to determine partnership in register-based census

Helle Visk
Chief Methodologist
Statistics Estonia

Abstract

In the preparation of its first register-based census in 2021, Estonia faces a challenge of dwelling-based household formation with inaccurate place of residence data. Pilot census in 2016 overestimated the number of lone parents by 67%. To increase accuracy, an index to measure probability of partnership is developed. Data on signs of partnership has been collected (marriages, divorces, mutual children, co-ownership of property *etc.*). In this paper, the partnership index is based on a logistic regression model. Model parameters are estimated and its accuracy is tested on partnership data from Estonian Labour Force Survey and Estonian Social Survey. To avoid putting the same person in multiple households, the partnership data is considered as a bipartite graph and stable marriage matching is calculated with Gale-Shapley algorithm. The model accuracy is 86%, the work on improving the index continues.