

DATA-DRIVEN HOUSEHOLD'S ARCHETYPES TO CLASSIFY BUILDING'S OWNER BUDGET RESTRICTIONS ON INVESTING IN DEEP RENOVATION

Iná Maia, TU Wien, maia@eeg.tuwien.ac.at
Lukas Kranzl, TU Wien, kranzl@eeg.tuwien.ac.at
Andreas Müller, e-think, mueller@e-think.ac.at

Overview

EU's deep renovation rates have not been sufficiently fast enough to significantly reduce building stocks' emissions. As a reaction against these facts, the newly published policy packages "Fit to 55" aim to boost renovation activities and to decarbonise the building stock until 2050. Nevertheless, building owners' affordability to pay for renovation has been recognised as a significant barrier.). In this paper, the authors follow a holistic approach by clustering the households according to techno-socio-economic characteristics, which is the innovative aspect of the present paper. Following research questions will be answered: 1) which variables available in the SILC and HBS databases can be used to define techno-socio-economic building types?; 2) For different techno-socio-economic building types can we observe different savings ranges?; and 3) By clustering the household according to the techno-socio-economic characteristics, which patterns can be observed?

Methods

The present paper contributes to this context by classifying households according to their budget restrictions, relevant information to develop more user-targeted financing instruments. Budget restrictions are expressed through the household's savings (income minus expenditures). The method follows two steps. First, a statistical matching using logistic regression was performed to merge EU-SILC and HBS datasets. The main objective of this step was to set up a techno-socio-economic dataset. Then, in the second step, a machine learning algorithm for clustering was applied to derive the households types according to techno-socio-economic characteristics. . The whole approach was carried out, tested and validated for data from Spain, because of the high number of observations in the database, consistency between the datasets and their completeness.

Results

First results presented a logistic regression model accuracy of 77% , meaning that the datasets could be successfully merged. Then, following variables were identified to derive techno-socio-economic household types: dwelling type, degree of urbanisation, tenure status, age, family composition, disposable income, total housing costs, and total household expenditure. Further results will identify patterns of the the households savings for the different household types.

Conclusions

Identifying patterns according to different household types are an important instrument to better design adequate financing schemes. In this context, end-user targeted financing schemes should help boosting building renovation rates, therefore achieving the EU's building stock decarbonisation goals.

References

- D'Orazio, Marcello, Marco Di Zio, and Mauro Scanu. 2006. *Statistical Matching: Theory and Practice*.
Donatiello, Gabriella, Marcello D. ' . Orazio, Doriana Frattarola, Antony Rizzi, Mauro Scanu, and Mattia Spaziani. 2014. "Statistical Matching of Income and Consumption Expenditures." *International Journal of Economic Sciences* (3).
Eurostat., European Commission. 2020. "European Union Statistics on Income and Living Conditions (EU-SILC)."
Eurostat., European Commission. 2021. "Household Budget Survey - Access to Microdata - Eurostat." Retrieved April 2, 2022 (<https://ec.europa.eu/eurostat/web/microdata/household-budget-survey>).
Eurostat. 2020. *Household Budget Survey Wavev - EU Quality Report (Version I)*.
Filippidou, Faidra, Maria Kottari, Savvas Politis, and Christiana Papapostolou. 2019. "Mapping Energy Poverty in the EU: Policies, Metrics and Data." *Eceee Summer Study Proceedings* 2019-June(June):1199–1207.

