Underlying determinants of energy efficiency and renewable energy adoption: factor analysis based on the Dutch household survey

Aktive Endkunden-/Prosumerpartizipation & Gebäudesektor

Ardak AKHATOVA[[1]](#footnote-1)(1), Erkinai DERKENBAEVA(2), Eveline VAN LEEUWEN(2)

and Lukas KRANZL

(1)TU Wien (2)Wageningen University

Motivation and Research Question

Understanding the main factors behind the household investments in energy efficiency retrofits (EER) and renewable energy (RE) adoption is essential for designing efficient and effective policies. In the Netherland, a number of factors have been identified as drivers and barriers of EER and RE adoption based on the earlier national household surveys [1]–[3]. However, the results of the regression analyses on the questionnaire survey suffer from high dimensionality (i.e. too many variables) and inability to explain the relationship between variables. To address these shortcomings, we conduct the exploratory factor analysis (FA) to identify the underlying factors behind a number of variables important for EER and EE adoption.

Data and Method

The data from Dutch national household survey (WoOn) from 2021 consisting of about 25,000 observations are taken as the basis for the FA. Socio-demographic, motivational, dwelling-related, economic and spatial variables that are found to be significant in the literature are selected and recoded into a suitable form (i.e. only numerical, binary and ordinal variables are usable in the FA). These variables are shown in Table 1.

Factor analysis is a statistical dimension reduction technique originating from psychology, namely, psychometrics. It assumes that multiple observed variables have similar patterns because they are associated with a latent, i.e. unobservable factor. The factor model is represented as:

$$Y = μ + αΦ +ϵ$$

where $Y=(y\_{1}, y\_{2}, …y\_{p})^{T} $is a vector of observable random variables, $μ=(μ\_{1},μ\_{2}…μ\_{p})$ is a vector of means, $α$ is a p × m matrix of loadings, $Φ=(ϕ\_{1},ϕ\_{2}, …ϕ\_{m})^{T}$ is a vector of common factors and $ϵ=(ϵ\_{1}, ϵ\_{2}, …ϵ\_{p})^{T}$ is a vector of residual random errors [4]. Here, $p$ represents the number of variables and *m* represents the number of common factors. By calculating the covariance matrix for these variables and running the factor analysis in R using the psych package [5], we interpret the factor loadings ($α$) and observe which variables are correlated with latent factors.

Results and Discussion

First of all, suitability of FA for the given dataset is defined using a standard Kaiser-Meyer-Olkin factor adequacy test [6]. The overall measure of sampling adequacy (MSA) for the given dataset is 0.73. According to [6], values higher than 0.5 indicate that it is reasonable to use FA for the given dataset.

Another important step is determining the optimal number of factors to be extracted. The parallel analysis does that by comparing the eigenvalues of the observed data to those of the random correlation matrices (i.e. generated via Monte Carlo simulation) [7]. This test showed that five factors seem to be reasonable. Factor analysis model is built using a mixed correlation matrix instead of a purely linear Pearson correlation matrix (because many of the variables are binary or ordinal) [5].

The results of the factor analysis with 5 factors using Principal Axis Factoring for factor extraction and the varimax rotation indicate that the first factor (PA1) groups neighbourhood-related factors. That is, dwelling type (1- apartment, 2- terraced, detached and semi-detached houses) is correlated with the urbanity of neighbourhood, social cohesion and insecurity). The second factor (PA3) groups economic variables and energy consumption. The third factor (PA2) seems to combine socio-demographic factors (age, education, household size and length of residence) and the willingness to adopt more EER and RE. PA5 shows a high correlation between whether a dwelling is a flat or apartment and whether they have solar panels. Finally, PA4 shows that older houses bought earlier groups with the higher gas consumption and absence of heat pump.

Table 1. Factor analysis output

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variables** | **PA1** | **PA3** | **PA2** | **PA5** | **PA4** |
| year\_bought | -0,122 | -0,017 | 0,620 | 0,189 | -0,413 |
| constr\_year | 0,084 | -0,113 | -0,006 | -0,061 | -0,386 |
| income | 0,021 | 0,509 | 0,211 | -0,053 | -0,003 |
| house\_value | 0,040 | 0,795 | 0,034 | -0,069 | -0,054 |
| wealth | 0,020 | 0,603 | -0,139 | -0,006 | 0,012 |
| usable\_area | 0,312 | 0,502 | 0,007 | -0,095 | 0,041 |
| elect\_cons | 0,351 | 0,428 | 0,209 | 0,019 | 0,040 |
| gas\_cons | 0,395 | 0,434 | -0,015 | -0,033 | 0,399 |
| soc\_cohesion | 0,849 | 0,166 | -0,077 | -0,092 | -0,138 |
| insecurity | 0,737 | 0,114 | -0,083 | -0,108 | -0,071 |
| age | 0,059 | 0,127 | -0,835 | -0,096 | 0,150 |
| education | -0,144 | 0,223 | 0,394 | -0,036 | -0,084 |
| hh\_size | 0,230 | 0,179 | 0,568 | -0,165 | 0,022 |
| willing\_inv\_ord | -0,066 | 0,105 | 0,408 | -0,106 | 0,004 |
| contact\_neigh | 0,094 | 0,033 | -0,069 | -0,083 | 0,015 |
| want\_to\_move | -0,154 | -0,076 | 0,151 | 0,118 | 0,068 |
| urbanity\_buurt | -0,809 | 0,023 | 0,085 | 0,097 | 0,105 |
| dwelling\_type | 0,697 | 0,084 | 0,172 | -0,827 | 0,283 |
| past\_out\_renov | -0,043 | -0,047 | -0,068 | 0,109 | -0,316 |
| past\_ind\_renov | 0,029 | 0,054 | -0,361 | 0,050 | -0,154 |
| exist\_insulation | -0,208 | -0,112 | -0,126 | 0,294 | 0,351 |
| exist\_double\_glazing | -0,041 | -0,018 | -0,099 | 0,161 | 0,234 |
| exist\_PV | -0,109 | -0,058 | -0,095 | 0,996 | 0,269 |
| exist\_heat\_pumps | -0,110 | -0,348 | -0,057 | 0,185 | 0,666 |
| Proportion of variance explained  | 0.12  | 0.09  | 0.09  | 0.08  | 0.06 |
| Cumulative Var  | 0.12  | 0.22  | 0.30  | 0.39  | 0.45 |

The results of the factor analysis show that indeed some variables may have underlying factors, which helps to reduce the number of variables in such analyses. To identify which of these factors are associated with each the decisions – insulation, double-glazing of windows, PV adoption and heat pump purchase, factors scores will be created and used for regression. The identification of the key factors will assist the integration of social aspects into energy models (e.g. agent-based models).

References

[1] S. Halleck Vega, E. van Leeuwen, and N. van Twillert, “Uptake of residential energy efficiency measures and renewable energy: Do spatial factors matter?,” *Energy Policy*, vol. 160, p. 112659, Jan. 2022, doi: 10.1016/j.enpol.2021.112659.

[2] S. Ebrahimigharehbaghi, Q. K. Qian, F. M. Meijer, and H. J. Visscher, “Unravelling Dutch homeowners’ behaviour towards energy efficiency renovations: What drives and hinders their decision-making?,” *Energy Policy*, vol. 129, pp. 546–561, Jun. 2019, doi: 10.1016/j.enpol.2019.02.046.

[3] S. Ebrahimigharehbaghi, Q. K. Qian, G. de Vries, and H. J. Visscher, “Identification of the behavioural factors in the decision-making processes of the energy efficiency renovations: Dutch homeowners,” *Build. Res. Inf.*, pp. 1–25, Jun. 2021, doi: 10.1080/09613218.2021.1929808.

[4] A. Basilevsky, *Statistical Factor Analysis and Related Methods*. Winnipeg, Manitoba: John Wiley & Sons, Inc., 1994.

[5] W. Revelle, “psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 2.2.9, https://CRAN.R-project.org/package=psych.” Northwestern University, Evanston, Illinois, 2022.

[6] H. F. Kaiser, “An index of factorial simplicity,” *Psychometrika*, vol. 39, pp. 31–36, 1974.

[7] J. L. Horn, “A rationale and test for the number of factors in factor analysis,” *Psychom. 1965 302*, vol. 30, no. 2, pp. 179–185, Jun. 1965, doi: 10.1007/BF02289447.

1. Energy Economics Group (EEG), Gusshausstrasse 25/CF0451, A - 1040 Vienna, Austria; +43 (0) 1 58801 370336, akhatova@eeg.tuwien.ac.at [↑](#footnote-ref-1)