Maciej Beręsewicz, PhD
Poznań University of Economics and Business
Department of Statistics
Poland
maciej.beresewicz@ue.poznan.pl

Cover Letter April 2, 2025

Dear Editors of the Journal of Statistical Software,

we are writing to submit an original manuscript of our article entitled ***nonprobsvy*** – *An R package for modern methods for non-probability surveys* for review to the Journal of Statistical Software.

Official statistics traditionally rely on probability surveys, censuses, or administrative registers that cover the entire population. However, rising non-response rates and survey costs have increased interest in non-probability data sources such as opt-in web panels, social media, scanner data, mobile phone data, and voluntary registers. Since these sources lack known selection mechanisms, standard design-based inference methods cannot be directly applied. That is why we have developed the `nonprobsvy` package in the R language (version 0.2.0; available on CRAN).

The `nonprobsvy` package has several advantages over packages currently available in R or Python. In particular, the novelty and our contribution can be summarised as follows:

- the package implements state-of-the-art methods recently proposed in the literature, along with valid statistical inference procedures (i.e. analytical and boostrap variance estimators),
- the package implements various approaches, such as calibrated inverse probability weighting, mass imputation, and doubly robust estimators, with our contributions that extend existing literature,
- the package supports the functions included in the `survey` package to account for the design of the probability sample (if is available).

We provide a user-friendly API that mimics `glm`, `svydesign` and other functions known in R, together with the main function to specify the approach and estimators.

The package has been developed since 2022 and the full history can be found at the GitHub repository `https://github.com/ncn-foreigners/nonprobsvy`. The package has been cited multiple times (cf. `https://scholar.google.com/scholar?q=nonprobsvy`), including in the review paper by *Cobo et al. (2024). Software review for inference with non-probability surveys. The Survey Statistician, 90, 40-47.* Our package is also included in *West et al. (2025). Applied Survey Data Analysis (3rd ed.). Taylor & Francis.*

As far as we know, the `nonprobsvy` is the only software (open-access or commercial) that offers such functionalities. That is why we believe that the paper and software will be of interest to the readership of the Journal of Statistical Software.

Thank you for your consideration of this manuscript.

Sincerely,

Maciej Beręsewicz, Łukasz Chrostowski & Piotr Chlebicki