

International Journal of Population Data Science

Journal Website: www.ijpds.org



Reproducibility and robustness of research conducted with administrative data: lessons learnt from the CBS Replication Games

Lenka Fiala¹, Jack Fitzgerald², Kasia Karpinska³, and Angelica Maria Maineri³

¹Institute for Replication (University of Ottawa), Ottawa, Canada; Tilburg University, Tilburg, Netherlands

²VU Amsterdam, Amsterdam, Netherlands

³Erasmus School of Social and Behavioral Sciences (ESSB), Rotterdam, Netherlands

The social sciences are experiencing a renaissance in open science, with mass replication initiatives and transparency policies becoming widespread in many institutions and academic journals. However, the rising use of restricted data is a key barrier to replication in the social sciences. In a sample of 158 articles published in Top 5 economics journals from 2020–2023, over 48% required restricted-use data to produce some main findings (Fitzgerald 2025). The proportions of papers that use restricted data have tripled or quadrupled in economic subdisciplines since 2000 (Garg & Fetzer 2024). The time and costs associated with re-accessing restricted-access data tend to be a significant barrier to preclude reproducibility in practice. In February 2026, ODISSEI and the Institute for Replication will host a Replication Games hackathon where teams of replicators will attempt to reproduce and run robustness checks

on six papers published since 2020 in top economics journals with robust transparency policies that are produced entirely from Statistic Netherlands administrative microdata and other public-use data, and have made code available. The results of our reproducibility and robustness analyses can help improve the trustworthiness of scientific analyses using restricted-access data. Difficulties in reproducing findings in these papers can help highlight necessary infrastructural adjustments (e.g., improvements in version control) to help ensure that findings using CBS microdata are reproducible. Conversely, if results in these papers are reasonably reproducible and robust, this can help boost trust in research using restricted-access data. The presentation will outline the preliminary results and lessons learnt from the three-day event.

