

International Journal of Population Data Science

Journal Website: www.ijpds.org



Tackling undercoverage in educational attainment registers using large-scale survey data: Statistics Netherlands' Educational Attainment File

Christian Fang¹, Karen van Hedel¹, Bram Hogendoorn¹, and Johan van Rooijen¹

¹Statistics Netherlands, The Hague, Netherlands

Educational attainment is a key variable in social-scientific research. In the Netherlands, most information on educational attainment is obtained from various registers of the Ministry of Education. These cover student enrollments in nearly all Dutch education institutions from the 1980s onwards. However, some residents have attained (additional) education before these registers came into existence, abroad, or at institutions not included in these registers (e.g. private institutions). This means that for 29% of the population information on education is entirely absent, whereas for some of the observed 71%, educational attainment is underestimated. Here, we present a unique method developed to deal with these deficiencies. By combining and integrating administrative data

and data from the Dutch Labour Force Survey in a unique way, Statistics Netherlands can make valid inference to the population level on the basis of incomplete coverage. We compute weights for individuals not observed in any register but who were observed in the survey, with these survey observations effectively being treated as though they are a random sample from the unobserved population. Furthermore, we impute scores for individuals observed in only registers to correct for the undercoverage of education. Finally, a bootstrap approach is applied to assess uncertainty of statistical outcomes. Taken together, this method allows valid inference to the population level, generating point estimates as well as estimates of uncertainty.

