Estimating the Validity of Administrative Variables

Bart F.M. Bakker

Statistics Netherlands, bfm.bakker@cbs.nl

Abstract: Administrative data have become more important for both official statistics and academic research. One possible problem with such data is that they are biased and have a low validity. Although this problem is often mentioned in a qualitative respect, the validity is seldom quantitatively measured. This article presents a method to estimate the validity of administrative variables. By applying the classical test theory, the validity can be determined by using linked survey and administrative data which should measure the same concepts. This idea is elaborated with an empirical example in which the construct validity of age, gender, educational attainment and wages is determined simultaneously. A linear structural equations model with a measurement component is used to compute the construct validity. The analyses reveal that educational attainment and wages show some bias, but not higher than the bias found in the survey.