

IAB-DiskAB

Einladung
15/2009

The Impact of Cleansing Procedures for Overlaps on Estimation Results – Evidence for German Administrative Data.

Patrycja Scioch (FDZ)

Process-generated and administrative datasets have become increasingly important for labor market research over the past ten years. Major advantages of this data are large sample sizes, absence of retrospective gaps and unit non-responses. Nevertheless, the quality and validity of the information remains unclear and a lot of preparation and data cleansing is necessary before the data are in shape to be evaluated. Unfortunately only few researchers provide access to their cleansing procedures and therefore also the impact of them on the results of the analyses is unidentified. This paper contributes to this subject and focuses on the variation of research results due to alternative data cleansing procedures. In particular, the paper replicates an evaluation study presented by Wunsch/Lechner (2008) and investigates the robustness of the results. Variance is induced by developing different cleansing procedures for overlapping and parallel observations. The descriptive results show that the differences between the data sets (based on the different procedures) are statistically not significant but show different magnitudes on some attributes. Similar results are to be found for the subsequent analysis of the treatment effects. The results emphasize that empirical findings seem to be robust to variations in the underlying cleansing procedure.

Dienstag, 3. November 2009

10.00 Uhr

Sitzungssaal 126a