

# Semiparametric efficient G-estimation with invalid instrumental variables

## ABSTRACT

The instrumental variable method is widely used in the health and social sciences for identification and estimation of causal effects in the presence of potentially unmeasured confounding. In order to improve efficiency, multiple instruments are routinely used, leading to concerns about bias due to possible violation of the instrumental variable assumptions. To address this concern, we introduce a new class of g-estimators that are guaranteed to remain consistent for the causal effect of interest provided that a set of at least  $\gamma$  out of  $K$  candidate instrumental variables are valid, for some  $\gamma \leq K$  set by the analyst ex-ante, without necessarily knowing the identities of the valid and invalid instruments. We provide formal semiparametric efficiency theory supporting our results. Both simulation studies and applications to the UK Biobank data demonstrate the superior empirical performance of our estimators compared to competing methods.