Double Boosting GMM for High Dimensional IV Regression Models[*]

Tae-Hwy Lee[†] and Hao Xu[‡]

March 2018

Abstract

When endogenous variables are approximated by sieve functions of observable instruments the number of instruments increases rapidly and many may be invalid or irrelevant. We introduce Double Boosting (DB) which consistently selects only valid 'and' relevant instruments. Also we show DB does not select weakly relevant or weakly valid instruments. DB works even when there are more instruments than the sample size. Monte Carlo compares DB-GMM with other methods such as GMM using Lasso. In the application of estimating the BLP-type automobile demand function, where price is endogenous and instruments are high dimensional functions of product characteristics, we find the DB-GMM estimator of the price elasticity of demand is more elastic than other estimators.

Key Words: Causal inference with high dimensional instruments, Irrelevant instruments, Invalid instruments, Weakly relevant instruments, Weakly valid instruments, Boosting, Price elasticity of automobile demand.

JEL Classification: C1, C5, D2, L6

[†]Department of Economics, University of California, Riverside, CA 92521. E-mail: taelee@ucr.edu
[‡]Department of Economics, University of California, Riverside, CA 92521. E-mail: hao.xu@email.ucr.edu