Endogeneity in Large Panel Models: Identification without (External) IV

Andrei Zeleneev

Abstract :
We present a novel approach to causal identification and estimation with large N and T panel data without external instruments. Existing approaches aim to extract and purge individual and/or period-specific factors from the outcomes and continuous treatments. They assume that the remaining idiosyncratic variation in treatment is exogenous. By contrast, we exploit the presence of an exogenous period-specific factor or factors (internal instruments) to identify causal effects. Thus, our approach is applicable when idiosyncratic variation in the continuous treatments (sources of variation that vary both over time and between individuals) is endogenous. Aggregate shocks (i.e., factors that vary over time but not between individuals) are commonly used to identify causal effects, but existing studies require that these shocks are directly observed, using them as external instruments. Our analysis demonstrates that causal effects are identifiable in such settings even if these aggregate shocks are unobserved. Our results allow for the presence of both endogenous and exogenous period-specific factors, and indeed the latter may only be exogenous after controlling for the former. Our identification results extend to settings with heterogeneous causal effects and to non-parametric additive models with heterogeneity. We propose a practical estimation method and provide simulation evidence of its efficacy.