

Partially Identified Treatment Effects for Generalizability

Wendy Chan

Will this intervention work for me? This is one of the questions that make up the core of generalization research. Generalizations focus on the extent to which the findings of a study apply to people in a different context, in a different time period, or in a different study altogether. In education, one common type of generalization involves examining whether the results of an experiment (e.g., the estimated effect of an intervention) apply to a larger group of people, or a population.



Figure 1. *The Role of Assumptions in Estimation*

Ideally, to generalize experimental results to a population, two tools are needed: (1) probability or random sampling to select which individuals or sites are part of a study, and (2) randomization to determine which individuals or sites in a study are offered an intervention. Random sampling ensures that individuals in the study are similar or representative of individuals in the population. Intervention randomization ensures that the study results are not sensitive to hidden biases due to systematic differences among individuals.

The challenge in educational studies is that random sampling is not common. As a result, when researchers use the data to generalize findings, assumptions are needed. Figure 1 illustrates

that without random sampling, assumptions are an important intermediate step between the data and the population estimates.

However, assumptions can be wrong. Or they can be strong in the sense that they make certain conjectures on the data, which cannot be empirically verified. The paper examines alternative ways to estimate the results without having to rely on strong assumptions that cannot be validated by the data. An important tradeoff of these methods is that instead of providing a single number, these methods provide a range of values, or bounds, that are plausible and consistent with the data. Because the bounds are estimated with few to no assumptions, they may be more credible in practice, especially in studies where the stronger assumptions are not met.

One of the limitations of bounds is that they can be too wide. The paper explores alternative assumptions that make fewer restrictions on the data to narrow bounds. As an example, one alternative assumption is that when given the option of two interventions, schools will choose the one that yields the most benefit for its students. While these alternative assumptions are still assumptions, they may be more plausible and credible in a study.

The paper includes an empirical example based on an experimental study that evaluated the effect a benchmark assessment system on students' test scores. Chan explains why the alternative assumptions are plausible for the study and provides the estimated bounds under each set of assumptions. In practice, assumptions play an important role when random sampling, treatment randomization, or both are missing. It is crucial for researchers to identify assumptions that are plausible since they affect the validity of the results.