MSI Webinar:

Part 2 – MSI Webinar: Incrementality: Methods for Causal Inference

December 8, 2022 | Virtual | 12:00 PM – 1:00 PM EST

Speakers:

Kathleen Li - *University of Texas at Austin*
Yanwen Wang - *The University of British Columbia*

Overview:

In part 2 of the MSI Series on Causal Measurement of Advertising Effects, speakers Yanwen Wang (The University of British Columbia) and Kathleen Li (University of Texas at Austin) detailed how marketers can examine causal relationships after treatments such as an ad or marketing intervention, by leveraging random controlled trials (RTC) or data gleaned from quasi-experiments. In the opening, Wang defined causal inference as "inferring the effect of one thing on another." She noted that from an econometrics perspective, there is an interest in the causal relationship between X and Y, but not an interest in understanding why the two are related. She indicated that the issue of causal inference arises when researchers want to fill in missing data using a variety of methods.

Wang indicated that in RCTs, the gold standard in demonstrating causality, a portion of units or groups are randomly chosen to be in the treatment group, whereas others are chosen to be in the control group. Using this method allows for the conclusion that there are no systematic differences in any of the variables (observables or unobservables). Results from this technique are very effective in comparing the average difference between the control group and the treated, to detect effects stemming from the treatment such as a marketing or ad intervention. Wang noted that there are situations where running an RCT is not an option. In these cases, she suggested using quasi-experimental techniques. Using this type of methodology can be effective by employing the use of observational data rather than running a field experiment. By doing so, marketers can still make a causal inference that includes observing data change after an exogenous variation, or shock (i.e., a change in the environment). Wang highlighted a variety of tools that can be leveraged in quasi-experiments to help fill in the missing group data.

In the second half of the presentation, Kathleen Li examined quasi-experiments and tools to leverage when carrying out this methodology in more detail. She considered techniques such as Difference-in-Differences, Synthetic Control, Matching and Random Forests (heterogeneous treatment effects) as methods to fill in the missing information. In her presentation, Li indicated that using Difference-in-Difference is the most popular of the techniques that mimic an experiment when the luxury of using randomization (RTC) is not possible. Another method considered when implementing a quasi-experiment is the Synthetic Control Method, which Li touted as a more "flexible version of the Difference-in-Difference method." This method can be applied in situations where a treatment isn't clear or known, through a "weighted average" to create an estimated version of the treatment (synthetic control). Leveraging Matching in a quasi-experiment can be employed by comparing similar units in the treatment and control groups. Finally, Li discussed the use of Random Forests as a technique to use when understanding heterogeneous treatment effects or how treatment effects "differ across individuals or groups."

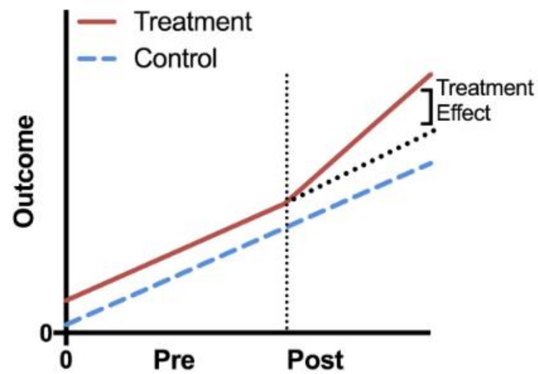
Takeaways

- The basic definition of **causal inference** is **"inferring the effect of one thing on another"** (i.e., did X cause Y?). **From an econometric perspective, there is an interest in the causal relationship between X and Y**, but not an interest in understanding why the two are related.
- The use of **causal inference stems from a missing data problem**, where we need to estimate the counterfactual. This requires researchers to make certain assumptions about the data.
- The gold standard in addressing the challenge of "missing data" is to use **Random Control Trials (RCT),** an experiment where some of the units or groups will be randomly chosen to be in the "treatment group," while others will be in the "control group" in order to detect differences between the groups that can be attributed to the treatment.
    - A treatment in this case can refer to an advertising or marketing intervention.
- In cases where an RCT is not feasible, a **quasi-experiment** can be used by employing observational data. **Quasi-experiments include observing data change after an exogenous variation, or shock** (i.e. a change in the environment).
    - Tools for conducting research with quasi-experiments include Difference-in-Differences, synthetic control, matching and random forests (heterogeneous treatment effects).

- A commonly used quasi-experiment approach is **Difference-in-Difference**, which can demonstrate change after a shock is applied. Assuming that trends in the treatment and control group would have been the same without the treatment, the results can be estimated using linear regression.

Difference-in-Differences (DID)

- Most popular quasi-experimental method
- Generally used in settings with many treatment and many control units



$$ATE_{DID} = (\bar{Y}_{Treat,Post} - \bar{Y}_{Control,Post}) - (\bar{Y}_{Treat,Pre} - \bar{Y}_{Control,Pre})$$

- When DID is not feasible, the **Synthetic Control Method** is another quasi-experimental approach that uses a weighted average of control units to create a synthetic version of the treatment unit.
- Another quasi-experimental approach is **Matching**. This technique can be leveraged by comparing similar units in the treatment and control groups. For example, researchers can match customers who are treated (visited the store) with similar customers who shop online.
- Researchers who may be interested in how treatment effects "differ across individuals or groups" (heterogeneous treatment effects) can employ the **Random Forests** technique, which requires a very large number of treatment and control units and covariates for effectiveness.

- **The below graphic from the presentation acts as a guide** to help researchers decide which method is best to apply **when conducting a quasi-experiment.**