Importance of Experimentation

1. Experiment: an intervention that creates variation in order to teach us causal questions. Deliberately create variation, but in a way that is not related to other features.
   1. Why is it important to experiment? For causal determination. Randomized controlled trials are the golden standard. Results from a well-constructed randomized experiment, even if small, contains information that even the largest, most dynamic observational data cannot.
      1. Vocabulary
         1. Treatment group: a group of subjects provided a new drug to determine efficacy, but compared to the control group.
         2. Control group: a group of subjects provided with no new drug, compared to the treatment group.
      2. Decisions should be concerned with counterfactual. We observe the state of the world if we have done *X*. What would we have observed if we have done *not X*, or *Y*. Difficult because we cannot observe both of these states of the world.
         1. Arguments based on intuition and anecdotes provide no solid ground for disputation and resolution.
         2. One should be skeptical of causal inferences based on observational (i.e., non-experimental) data because of the possibility of unobserved heterogeneity.
   2. Intervention is the key element of an experiment. Experiments do not have to involve randomization, and randomization may be difficult to implement. We learn more and have better guarantees with deliberate variation than without it.
      1. Without randomization, there may be confounding factors which also influences the outcomes. We can control for confounds by repeating the experiment many times.
      2. Randomization produces the guarantee of independence.
      3. Possible pitfalls in making casual inferences from observational data (e.g., Nurses’ Health Study vs. Women’s Health Initiative on HRT)
   3. Observational Data can lead to “Natural Experiments” (naturally occurring data that a researcher reasons has the same properties as a true, designed and controlled experiment)
      1. Herschel’s Garden: the idea of viewing the night sky as a garden that features the same types of plants at different stages of development. Harder to apply to social data sciences due to un-observable heterogeneity.
      2. Best data from naturally occurring experiments: situations where variation was produced by something like random assignment (e.g., charter school lottery deciding applicant acceptance)
      3. Less ideal data: decent reasons to believe that those who received an intervention are otherwise identical to those who did not (e.g., snow plowing occurs on alternating days)
   4. Magic on the Internet: using field experiments to test equivalence between auction formats. Observational data holds no guarantees that auction format is independent of card value.
      1. Purpose and Method
         1. Used early online auctions. Place bids via email. Cards sold/auction closed, if there were no increases in 3 days. Can have ascending-bid or sealed-bid format.
         2. Auction theory: strategies exist, but lab research shows violation to theory.
            1. Strategic equivalence: strong prediction; there is a dominant-strategy mechanism for truthful revelation of valuation. Optimal strategy is the same bid: bid my maximum willingness to pay.

Dutch (descending-clock) and first-price auctions are strategically equivalent because they have the same amount of information

English and second-price auctions are strategically equivalent under private value (like those assigned in the lab)

* + - * 1. General revenue equivalence: weaker prediction; expected revenue of all four formats should be the same if people are risk-neutral.
      1. Field experiments are similar to observational studies. Auctions for real goods, by real people, who are really accustomed to bidding for these goods. Similar to lab experiments (deliberate, designed interventions into a real system, without actors being aware of the interventions). Two different auction formats with the same good and the same bidders.
         1. Similar to the real world. Cannot control individuals’ valuations of the goods. Cannot control risk, time, or other preferences of the subjects.
    1. Results and Conclusions
       1. Treatment unit with card-level data: Violation of Dutch/first-price revenue equivalence. Dutch format returns more profits. No qualitative difference of treatment order. Opposite of the violation that had been observed in the lab.
       2. Treatment unit with bid-level data: violation of strategic equivalence. Compared bids by the same bidder in two matched auctions when possible. Higher bids in Dutch auctions.
          1. Data censoring in Dutch auctions: we don’t see most people’s revealed prices because there is only one bid that is made which closes the auction.
          2. English auction produces slightly more revenue than second-price auction, but limited statistical power.
       3. Dutch/First-price raised more revenue than English and second-price auctions.
       4. Revenue ranking: Dutch, first-price, English, second-price
       5. Field data can produce different results from laboratory data, or can replicate the lab results. Various reasons why that can happen.
  1. Concluding Thoughts
     1. Importance of interventions: holding secondary factors constant while trying alternate methods
     2. Try new, uncertain things. Be willing to live with short-term loss for longer-term gains.

Three key elements to an experiment:

* Randomization
* Excludability: the only reason why we are observing the treatment effect is only due to the experiment, not because of the randomization process
* Non-interference: the dosage will only affect the unit, not any other unit.