Causal Effects, Experiments, and Identification

Avi Goldfarb
University of Toronto

Workshop on Quantitative Marketing and Structural Econometrics 2013

Rotman
UNIVERSITY OF TORONTO

A Simple Linear Model

- Take the equation

$$y=X\beta+e$$

- Does x cause y?
- What x can be used to best predict y?
- Applying variants of this simple framework is perhaps the dominant empirical method in the economics of labor, health, public finance, and innovation.
- With the right data, it can be a powerful tool for answering a variety of important marketing questions.

Rotman

Quasi-Experimental Econometrics

- This is a slightly different ("descriptive") empirical tradition than the one described in the previous session.
- It leverages the idea that experiments can suggest causal relationships.
- Then it looks for settings that mimic the conditions of an experiment under the right assumptions.
- If derived from a model, it can be seen as a reduced form. For example, many papers present quasi-experiments as the reduced form of a production function.
- It is perhaps best exemplified in Angrist and Pischke's book "Mostly Harmless Econometrics" or, on the technical side, Imbens and Wooldridge's JEL article "Recent Developments in the Econometrics of Program Evaluation"

Rotman

Examples in marketing

- Busse, Silva-Risso, and Zettelmeyer (2006 AER)
- Chevalier and Mayzlin (2006 JMR)
- Anderson, Fong, Simester, and Tucker (2010 JMR)
- Danaher, Dhanasobhon, Smith, Telang (2010 Mkting Sci)
- Choi and Bell (2011 JMR)
- Goldfarb and Tucker (2011 Mngmt Sci)
- Dhar and Bayliss (2011 JMR)
- Chen, Wang, and Zie (JMR 2011)
- Sun (2012 Mngmt Sci)
- Bronnenberg, Dube, and Gentzkow (2013 AER)
- Sun and Zhu (Forthcoming Mngmt Sci)
- Etc.

Rotman

Examples

- Busse, Silva-Risso, and Zettelmeyer (2006)
  - Do auto retailers pass through manufacturer promotions to customers?
- Chevalier and Mayzlin (2006)
  - Do online product reviews affect sales?
- Goldfarb and Tucker (2011)
  - Does privacy regulation reduce the effectiveness of online advertising?
- Sun and Zhu (Forthcoming)
  - Does advertising change the type of content offered by online media?

Rotman

Developing a great paper with quasi-experiments

- Research question
- Identification strategy
- Mechanism

Rotman

The Research Question

- Is y interesting?
- Is X interesting and under someone's control?
- What does "interesting" mean anyway?

Rotman

Your goals

- Find out what is in the data
  - What is the "experiment"?
  - Are there things that confound the experiment?
  - Can you show those things don't matter? If not, how and why do they matter?
- Interpret the data
  - What are the big picture consequences for a causal relationship?
  - Does this interpretation suggest additional tests?
  - What assumptions allow you to use this interpretation?
- Communicate what you found
  - Emphasize your core results
  - Be clear about the caveats—honesty is the best policy!

Rotman

Identification strategy

Rotman
UNIVERSITY OF TORONTO

Identification

- Why the obsession?

Rotman

Identification (from Heckman 2000)

- The problem of identification is that "many different theoretical models and hence many different causal interpretations may be consistent with the same data".
- "The econometric analysis of the identification problem clarifies the limits of purely empirical knowledge"
- "The justification for interpreting an empirical association causally hinges on the assumptions required to identify the causal parameters from the data"

Rotman

Identification

- For any discrete event/policy, each individual has two possible outcomes
  - Y_{i1} if the individual experiences the event
  - Y_{i0} if the individual does not experience the event
- The difference between the two is the causal effect.
- The identification problem is that only one outcome is observed for each individual because you can't both receive the treatment and not receive the treatment.
- The unobserved outcome is called the "counterfactual". The unobservability of the counterfactual means assumptions are required.

Rotman

What is "endogeneity"?

- Endogeneity means that those who experience the event and those who don't are different in some relevant unobserved way(s)
- The goal is to make that "unobserved way" as irrelevant as possible

Rotman

Understanding the unobserved outcome

- Random assignment solves this problem as the difference between the group that experiences the event and the group that doesn't is, by definition, independent of other factors.
- Therefore random assignment is often called the "gold standard" of identification.
- Often experiments are not feasible, not appropriate, or not worth the cost.
- In this case, the objective is to identify something that approximates random assignment

Rotman

Approximating Random Assignment

- So researchers look for exogenous variation: "Policy" changes that affect some groups and not others
- Policies can include variation by
  - Country, state, city, firm, establishment, street corner, individual, publication, invention, "act of God", website visit, behavior by others that won't care about the response (technology adoption, market entry, advertising, pricing,…), etc.
- Ideally, you observe before and after for the treatment and the same time period for the control

Rotman

Objectives in research design

1. Exogenous variation in the explanatory variables (i.e. the "treatment")
   - The mechanism for treatment assignment needs to be clearly understood. "Being able to rule out obvious sources of endogeneity is not enough" (Meyer 1995, p. 153)
2. Finding a control group that is comparable
3. Probing the implications of the hypothesis under the test

Rotman

Types of endogeneity

- Omitted variables
  - Other things may happen at the same time as the treatment
- Simultaneity (reverse causality)
  - The treatment may be affected by the apparent outcome
- Selection
  - The treated population may be unrepresentative
  - Often addressed with structure or new data, rather than with the toolkit I will emphasize in this session

Rotman

Tools for identification

- Controls (the old fashioned way!)
- The “diff-in-diff”
- Regression discontinuity
- Instruments
- Others: Matching, selection estimators, bounds.

Rotman

Controls

- Think back to when you first learned multiple regression
- The reason to add variables is explicitly to turn "omitted variables" into "variables"
- E.g. Wooldridge
  - "Multiple regression analysis is more amenable to ceteris paribus analysis because it allows us to explicitly control for many other factors that simultaneously affect the dependent variable"
  - "Because multiple regression models can accommodate many explanatory variables that may be correlated, **we can hope to infer causality in cases when simple regression would be misleading**"
- The problem is that you never know if you are capturing all the relevant omitted variables…

Rotman

Assessing the scope of the omitted variables problem

- Altonji, Elder, and Taber (2005) provide a method to compare the role of the included and omitted variables
- The idea is to examine how much the effect of interest changes as controls are added
- Then ask, how important would the omitted variables have to be for the treatment effect to go away
- For example:

	Full Sample: Controls				Catholic 8th Grade Attendees: Controls			
	None (1)	Family Background, City Size, and Region[a] (2)	Col. 2 Plus 8th Grade Tests (3)	Col. 3 Plus Other 8th Grade Measures[b] (4)	None (5)	Family Background, City Size, and Region[a] (6)	Col. 2 Plus 8th Grade Tests (7)	Col. 3 Plus Other 8th Grade Measures[b] (8)
	A. High School Graduation							
Probit	.97	.57	.48	.41	.99	.88	.95	1.27
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	[.123]	[.081]	[.068]	[.052]	[.105]	[.084]	[.081]	[.088]
Pseudo R^2	.01	.16	.21	.34	.11	.35	.44	.58

- They then provide details on how to more formally assess how big the effect of the omitted variables need to be relative to the controls.

Rotman

This is often enough!

- If you do not have reason to expect reverse causality
- And if you do not have reason to expect substantial selection bias
- And if a large number of reasonable controls do not change your estimated treatment effect
- Then clearly state the assumptions behind your interpretation and move to exploring the mechanism and/or exploring the broader consequences.

Rotman

Diff-in-diff

Rotman
UNIVERSITY OF TORONTO

What is "diff-in-diff"?

	TREATMENT	CONTROL
BEFORE	A	B
AFTER	C	D

(C-D)-(A-B)==the effect of the treatment on the treated

- In regression format, with fixed effects, it simplifies to:

$$Outcome_{it}=\beta TreatmentGroup_i x AfterTreatment_{it}+\mu_i+\nu_t+\varepsilon_{it}$$

- Then we add controls γX_{it} to address additional omitted variables concerns. Difference out the fixed effects to avoid the incidental parameters problem.
- Examples: policy changes (e.g. privacy or advertising bans), consumer migration, offline store openings

Rotman

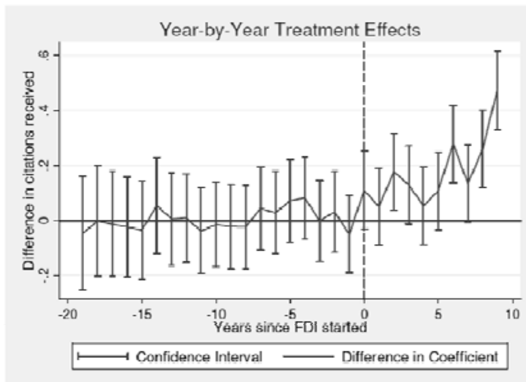
Diff-in-diff

- Meyer (1995) notes
  - "Good natural experiments are studies in which there is a transparent exogenous source of variation in the explanatory variables that determine treatment assignment"
  - "If one cannot experimentally control the variation one is using, one should understand its source"
  - Diff in diff is best when the control group before and after has a distribution of outcomes (Dependent variables) similar to the treatment group before.
    - Otherwise transformations of the dependent variable (e.g. using logs) may lead to different conclusions

Rotman

Communicating Pre-Trends

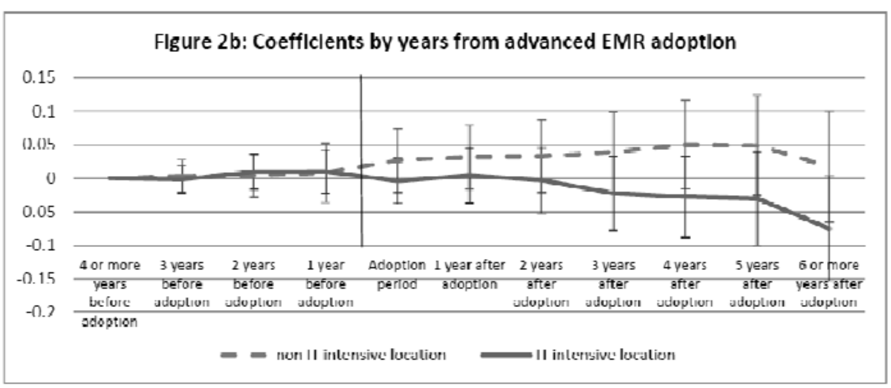
Figure 4b. Difference in average citations per type of bidder



Note: This figure divides the sample between winning and losing bidders and analyses the difference in the citations they receive over time, including confidence intervals at the 95% level. The zero value in the horizontal axis defines the year in which the privatisation case was resolved. We observe a number of things: (1) before FDI, there is no statistically significant difference in citations received by winners and losers; (2) after FDI, the gap gradually increases in favour of the winning bidder, reaching statistical significance a number of years after the privatisation case was resolved.

Rotman

Communicating Pre-Trends



Rotman

A Diff in Diff Etiquette

1. Explain and defend the "experiment"
2. Show treatment and control groups are similar pre-treatment
3. Compare/control for pre-treatment trends
4. Present the raw data
5. Present baseline estimates
6. With standard error corrections as appropriate
7. Robustness checks galore
8. Explore the mechanism
9. Apologize for all you can't do and give caveats

Rotman

Regression discontinuity

Rotman
UNIVERSITY OF TORONTO

Regression Discontinuity

- "The basic idea behind the RD design is that assignment to the treatment is determined, either completely or partly, by the value of the predictor being on either side of a fixed threshold"—Imbens and Lemieux (2008)
- The predictor may be correlated with outcomes, but the association is assumed to be smooth.
  - Therefore any discontinuity in the effect is assumed to be due to the treatment
- External validity is limited unless there is reason to assume homogenous treatment effects or unless the threshold population is inherently interesting
- Classic examples
  - Scoring policies for marketing offers ($49 vs. $51 monthly spend)
  - Government policies based on firm size
  - Time?

Rotman

Regression Discontinuity Etiquette

1. Explain and defend the "experiment"
   - Any discontinuity in the effect is assumed to be due to the treatment
   - The predictor may be correlated with outcomes but the association is assumed to be smooth
2. Show the treatment and control groups are similar
   - Similar to the graphs for diff-in-diff based panel results
3. Present the raw data
4. Present baseline estimates
5. With standard error corrections as appropriate
6. Robustness checks galore
7. Argue why the threshold population is inherently interesting or why treatment effects are likely to be homogenous
8. Apologize for all you can't do and give caveats

Rotman

Busse, Silva-Risso, and Zettelmeyer

$$(1) \quad P_{ijt} = \lambda_c CustCash_{jt} + \lambda_d DealCash_{jt} + \beta_1 X_i + \beta_2 X_{jt} + \beta_3 DealerComp_{ij} + \mu_j + \tau_{JT} + \varepsilon_{ijt}.$$

"the regression discontinuity approach dictates that we use data only immediately before and after a change in *customer cash* promotions but ***not*** data surrounding changes in dealer cash promotions."

TABLE 2—PRICE EFFECTS, BASIC RESULTS[†]

	Difference in differences	Regression discontinuity	
	(1)	(2A)	(2B)
Customer cash	−0.88	−0.81	−0.78
	(0.03)**	(0.07)**	(0.12)**
Dealer cash	0.39	0.38	0.31
	(0.07)**	(0.14)**	(0.07)**
GM Card	1.06	1.13	1.13
	(0.03)**	(0.10)**	(0.09)**
Competition	−7.66	−15.03	−18.63
	(5.60)	(9.59)	(9.63)*
Female	144.17	139.74	203.25
	(12.01)**	(45.20)**	(57.59)**
% Asian	−261.75	−188.40	−115.40

Rotman

Busse, Silva-Risso, and Zettelmeyer

- "**We now test the validity of the two key assumptions that were maintained when identifying these effects**."
- "**The identifying assumption of the difference-in-differences approach is** that the prices of cars in the same segment that are not on promotion in a given week are a valid counterfactual for the prices that would have been obtained on the promoted car in the absence of a promotion."
- "**Although we cannot observe this directly, we can examine** the trends of promoted and nonpromoted cars in the period just prior to the promotion. If the trends are similar between cars that are soon to be promoted and cars that are not, that gives some assurance that the nonpromoted cars are a valid counterfactual in the promotion period."
- **"The key maintained assumption in the regression discontinuity approach is** that transaction prices during the week just before the promotion starts are a valid counterfactual for transaction prices during the first week of a promotion."
- "**This would be violated if** the customers who purchase just before a promotion starts differed in some way that was related to negotiated prices from customers who purchase just after the promotion starts. In particular, this would be the case if there are "deal-prone" customers, who are particularly effective negotiators, and who wait to purchase a car until a promotion is offered. This would mean that the set of customers whom we observe buying before the promotion would pay higher prices on average, *with or without a promotion*, than the set of customers whom we observe buying during a promotion would pay, *with or without a promotion."*

Rotman

Instrumental variables

(Peter Rossi will cover this in depth tomorrow)

Rotman
UNIVERSITY OF TORONTO

Instruments in the experimental paradigm

- Instruments can be seen as natural experiments that affect the endogenous covariate indirectly.
- Regressing the instrument on the second-stage dependent variable will get you the result, but it will be scaled wrong:

  IV: $x=\gamma z+u \quad y=\beta\hat{x}+e$

  $\rightarrow y=\beta\gamma z+e$

- Therefore, you need two stages to get the elasticity right but the experiment happens at the level of the instrument and so the intuition on causality happens at the level of the relationship between z and y

Rotman

For completeness, An Instrumental Variables Etiquette

1. Explain and defend the "experiment"
   - You cannot actually ***test*** the validity of the exclusion restriction
2. Tests for the first stage correlation
   - Report the first stage and think about whether it makes sense
   - Report the F-statistic on the excluded instruments—are the instruments weak? How weak?
     - Note that IV is biased but consistent. Bias is large when instruments are weak.
   - Report the overidentification test if multiple instruments
   - Pick your best instrument and report just-identified results
     - Bias is less likely in the just-identified case
3. Report a Hausman test of whether the IV estimates are any different from OLS. Does the direction of the difference make sense?
4. Do a reduced-form regression of the dependent variable (2nd stage) on the instruments? Does it work?
   - This test rarely actually appears in the paper (but maybe it should...)
5. Robustness checks galore
6. Apologize for all you can't do and give caveats

Source: Angrist and Pischke (2009)

Rotman

Other tools

Rotman
UNIVERSITY OF TORONTO

Other tools

- Matching estimators
  - Rather than assuming the linear structure, matching estimators allow for a non-parametric (i.e. flexible) relationship for controlling observables.
  - If outcome measures are costly to obtain, matching saves time and effort in the data collection process
  - Matching estimators are still about controlling for unobservables
- Heckman correction/Selection estimators
  - For identification, need *instruments* that shift the selection probability but not the outcome
  - In the absence of strong instruments, the inverse Mills ratio terms are simply non-linear functions of the covariates and are only identified off functional form
- Bounds
  - Identifies what can be said without any assumptions, then add assumptions to narrow the bounds on the treatment effect
- Etc.

Rotman

Data structure

Rotman
UNIVERSITY OF TORONTO

Data structure

- A key decision is the choice of the unit of observation
- At what level does the dependent variable move?
  - Aggregate up to this level (at least)
- At what level does the main treatment variable move?
  - If you don't aggregate to this level, adjust your standard errors as appropriate (Donald and Lang; Bertrand, Duflo, and Mullainathan)
- At what levels do the controls move?
- Write out the estimating equations carefully. Pay close attention to the subscripts: They will help you determine if you have the data structure right.

Rotman

Developing a great paper with quasi-experiments

- Research question
- Identification strategy
- Mechanism

Rotman

Mechanism

Rotman
UNIVERSITY OF TORONTO

Falsification tests/Mechanism checks

- Falsification tests are about finding an example where your confounds would suggest the same result but your theory suggests otherwise.
- Essentially about understanding the setting and the underlying economics
  - What are the sources of non-randomness of assignment?
    - Omitted variables, selection, and/or simultaneity?
  - What other outcomes would be affected by these sources of bias that would not display the causal effect of interest?
  - What other groups would be affected by these sources of bias that would not display the causal effect of interest?
- If the effect goes away when theory suggests it should, then this helps identify the mechanism.
- Correspondingly, if the effect is larger when theory suggests it should be, then this also helps identify the mechanism.

Rotman

The value of the mechanism check

- Identifying an interesting main effect is typically just the first step.
- You also want to provide an understanding of the effect.
- This typically is about identifying heterogeneous treatment effects
- If the effect goes away when theory suggests it should, then you have likely identified the theory that drives the result
- So, after showing "privacy regulation hurt online advertising", we showed that it especially hurt unobtrusive advertising and advertising on general interest websites
- After showing "offline advertising bans increase online advertising effectiveness", we showed that it especially increased effectiveness for new and low awareness products.

Rotman

External validity

Rotman
UNIVERSITY OF TORONTO

A note of caution

- External validity matters
- The treatment may not be what you want to study
- The treated population may be unrepresentative
- In these cases, structure plays an important role. A model can enable you to use your estimates to say something about a counterfactual for which you lack explicit data.
- The treatment effect may be heterogeneous
  - Across places
  - Across institutions
  - Across time
  - Across demographics
    - (Though this can be an opportunity to identify mechanisms)

Rotman

Heterogeneous Treatment Effects

- Early studies assume that the effect of a treatment was constant
- This is unlikely to be true. This will affect what is actually learned from the analysis
- In other words, there are many kinds of treatment effects
  - ATE—Average treatment effect
    - What is the effect on all units?
  - ATT—Average treatment effect on the treated
    - What is the effect on the treated units?
    - If the treated units are different from many control units, this may be more interesting (though matching solves this too)
  - LATE—Local average treatment effect
    - What is the effect on the subpopulation that is induced by the treatment to change behavior?

Rotman

ATE, ATT, and LATE

- ATE≠ATT when treatment effects are heterogeneous
- LATE≠ATT and LATE≠ATE when not everyone is a "complier"
- Given a treatment, some people respond and others don't. There are three ways not to respond:
  1. Never-takers: Never take the treatment (always do 0)
  2. Always-takers: Always take the treatment (always do 1)
  3. Defy: Always do the opposite of the treatment assigned
- Take a random coupon drop by a store that regularly uses coupons. Some people never use coupons. Some people always use coupons. And some people may even be suspicious of the drop but seek coupons when they don't get them easily
- Under the assumption of no defiers, we can identify the LATE
- Then the size of these groups determines the match between LATE and ATE. Unfortunately, we cannot observe who is in which group.

Rotman

Summary

Rotman
UNIVERSITY OF TORONTO

Questions to ask when reading and writing quasi-experimental econometrics papers

- What is the research question?
- What are the core identification challenges?
  - Omitted variables? Selection? Simultaneity? In which way?
- What is the data structure?
  - For the dependent variable? For the treatment variable? For the controls?
  - What unit of observation is used in estimation?
- What is the core estimating equation?
  - How robust are the results to various identification checks?
- What is the main effect found?
  - Does the interpretation follow the data?
  - Is the research setting similar enough to the setting of broader interest?
- What is the mechanism identified?
  - Does the interpretation follow the data?
- How is the data communicated?
  - Are the caveats clear?

Rotman

A reading list

- Angrist, Joshua D. and Jörn-Steffen Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press: Princeton NJ.
- Meyer, B. (1995), "Natural and Quasi-Experiments in Economics," *Journal of Business and Economic Statistics*, 12, 151-162.
- Heckman, J. (2000), "Causal Parameters and Policy Analysis in Economics: A Twentieth Century Retrospective" *Quarterly Journal of Economics*, 115, 45-97.
- Altonji, J. G., T. Elder, and C. R. Taber (2005). Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools. *Journal of Political Economy* 113(1): 151-184.
- Moffitt, Robert (2005), Remarks on the Analysis of Causal Relationships in Population Research. *Demography* 42(1): 91-108.
- Bertrand, M., E. Duflo and S. Mullainathan (2004), "How Much Should We Trust Differences-in Differences Estimates?" *Quarterly Journal of Economics*, 119, 249-76.
- Donald, S. and K. Lang (2007), "Inference with Difference in Differences and Other Panel Data" *Review of Economics and Statistics*, 2, 221-233.
- Lancaster T. (2000) "The incidental parameter problem since 1948" *Journal of Econometrics*, 95: 391-413.
- Imbens, Guido W., and Thomas Lemieux (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142, 615-635.
- Manski, C. (2007), *Identification for Prediction and Decision*, Harvard University Press.
- Angrist and Kruger (2001) "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives*, 15, 69-85.
- Imbens, Guido W., and Jeffrey M. Wooldridge (2009). "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47(1): 5–86.
- Ai, Chunrong, and Edward C. Norton (2003) "Interaction Terms in Logit and Probit." *Economics Letters* 80, 123-129.
- McCloskey D, Ziliak S T. (1996). "The Standard Error of Regressions." *Journal of Economic Literature* 34(1): 97-114.
- Griliches, Zvi (1984). Data Problems in Econometrics. NBER Technical Working Paper #39.
- Manski, Charles (1993). "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, 60(3), 531-542.
- All articles. *Journal of Economic Perspectives*. Spring 2010 Symposium "Con out of Econometrics"

Rotman

Thank you

Rotman
UNIVERSITY OF TORONTO