

Investigating the Social World

Week 4





Welcome back!



Recap: Using “numbers” to investigate the social world

- ▶ Quantitative and qualitative social research can be very different.
 - Ontology/epistemology: **Positivist v constructivist**
 - Size of observations: **Large-N v small-N**
 - Nature of theoretical claims: **Theory as paradigm v theory as approach**
 - Nature of conclusion: **Probabilistic v deterministic**
 - Goal of inquiry: **Capturing the "universal" truth v delineating multiple "realities"**
- ▶ However, very often their differences may be misplaced or exaggerated.



	<i>A positivist approach</i>	<i>A constructivist approach</i>
What is ‘reality’?	A definable ‘reality’ or ‘truth’ exists and is observable	There is no ‘reality’ or ‘truth’ beyond our experiences
What is the goal of academic enquiry?	Acquisition of the ‘truth’	A more informed construction of the world
How are the researcher and the ‘researched’ related?	The researcher is independent of the ‘researched’	The researcher is not independent of the ‘researched’
What should be the role for values?	None - objectivity sought	Part of ‘reality’ - subjectivity celebrated
What kind of approach?	Predominantly based on observability or measurability and with the aim of seeking ‘evidence’	Predominantly based on discourse and meaning with the aim of seeking a more informed understanding of the world
What kind of data is preferred?	Predominantly quantitative	Traditionally associated with a predominantly qualitative approach
Examples of such studies in Development Studies	Dollar and Kraay (2002) <i>Growth is Good for the Poor</i>	Narayan <i>et al.</i> , (2002) <i>Voices of the Poor</i>



	Economics	Politics	Sociology	Social Anthropology
What is reality?	One reality exists; reality is <i>independent</i> of our thoughts; what is observable is real	One reality exists. reality is <i>independent</i> of our thoughts	One reality exists; reality is <i>independent</i> of our thoughts, but much of reality is unobservable	There are <i>different</i> realities associated with different standpoints and cultures
Goal of enquiry?	Acquisition of a <i>single</i> truth -- a universal, general law	We can establish truths or <i>generalizations</i> about human beings	Truth needs to be understood in terms of <i>practical adequacy</i>	<i>Interpretation</i> of local meanings; there is <i>no universal truth</i>
How are the researcher (you) and the "researched" related?	The researcher is <i>objective</i> and is <i>independent</i> of the "researched"	The researcher is <i>subjective</i> and is <i>not independent</i> of the "researched"	The researcher is <i>subjective</i> and is <i>not independent</i> of the "researched"	The researcher is <i>subjective</i> and is <i>not independent</i> of the "researched"



Sumner, A, and M Tribe. 2008. *International Development Studies*. London: Sage, p.72.



Recap: Using “numbers” to investigate the social world

- ▶ Quantitative social research focuses on producing systematic “inference” based on numeric data (Tufte 1974; KKV 1994).
 - **Descriptive:** Use basic statistics (e.g., mean, median, and standard deviations) to describe noticeable patterns in the data
 - **Statistical/causal:** Use (multiple) linear/non-linear regression to evaluate the relationship between **explanatory** (or “independent”) and **outcome** (or “dependent”) variables
- ▶ Our focus today is **causality**; namely, we will talk about why **causal inference** has been growing to dominate (quantitative) social research.



JQD:DM

Journal of Quantitative Description:
Digital Media

JQD

Following

Journal of Quantitative Description: Digital Media

@journalqd

Quantitative descriptive social science about digital media. No causal claims.

Open access. No fees.

Submit: journalqd.org/loi

mastodon.world/@journalqd

 journalqd.org  Joined August 2020

2 Following 2,245 Followers

<https://twitter.com/journalqd>



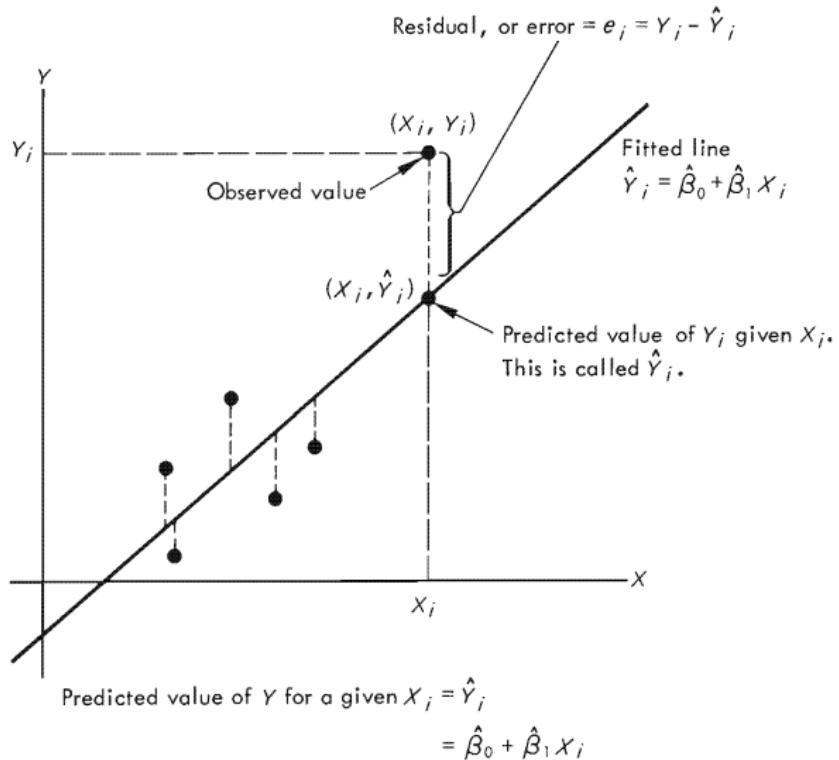


FIGURE 3-3 Notation for least-squares regression





Causality in social research



Why causality matters

- ▶ Social science research is often motivated by simple (and yet important and difficult) **cause-and-effect** questions. To name a few:
 - American politics: Does historical black slavery have a lasting impact on voters' attitudes toward the African American people?
 - Comparative politics: Does the descriptive representation of women and racial minorities in the legislature mitigate people's bias against them?
 - Political economy: Do cash-transfer programs reduce poverty?
 - International relations: Does international peacekeeping reduce conflicts? Does foreign aid work?
- ▶ Causality is in particular crucial for "evidence-based" policymaking.



Causality is hard to establish

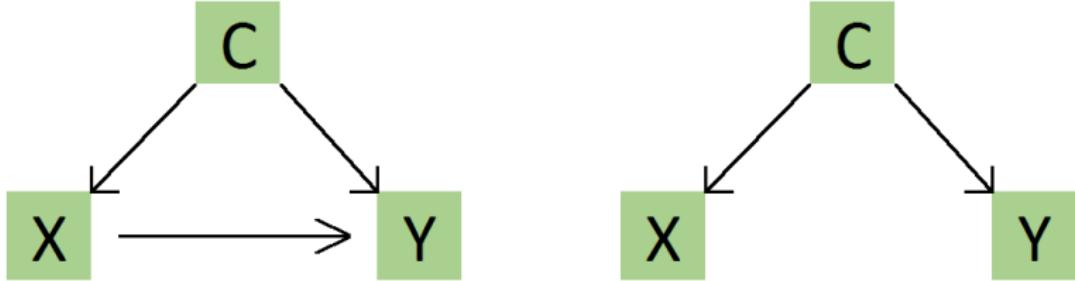
- ▶ Presence of confounders or common causes
- ▶ True counterfactual cannot be observed
- ▶ Causality is a complicated concept



Challenge 1: Correlation and common causes (confounders)

- ▶ Two variables X (independent or explanatory) and Y (dependent or outcome) can be **correlated** with each other under at least **three** conditions.
 - X causes Y .
 - X and Y have a common **cause**.
 - X and Y have a common **outcome**.
- ▶ To make sure X causes Y , we have to rule out the influence from all common **causes**, namely the **confounders**, through **statistical controls** in multiple regression.
- ▶ It is simply impossible and unrealistic to include all possible confounders in multiple regression.





Correlation with or without causation. In the directed acyclic graph (DAG), we use **arrows** to denote causality between two variables. Here X , Y , and C refer to **cause**, **outcome**, and **confounders** respectively.





- ▶ On average, neighborhoods that have more police officers also see more residents killed by gun violence.
- ▶ Is the police incompetent? Or even worse, are the police officers involved in some covert collusion with the local crime syndicate?





- ▶ On average, no statistically significant results exist to show that UN peacekeeping troops reduce the likelihood of communal violence.
- ▶ Is sending UN peacekeeping troops a waste of time and money?



Challenge 2: Counterfactual virtually does not exist

- ▶ Causal inference is difficult also because of the famous "fundamental problem of causal inference."
 - For any individual, we can never observe what actually happens under **counterfactual** in real life (Holland 1986).
 - Instead of comparing the difference at the **individual** level, we can compare the **average difference** in terms of the outcome of interest between **two (comparable) groups of people**.
- ▶ The comparison is not straightforward due to the presence of **confounders**.
 - Researchers usually attempt to **eliminate the influence of confounders on the cause** by randomly assigning our observations (e.g., survey respondents) into one of the experiment groups.
 - **Randomization may not be a perfect solution** (Deaton and Cartwright 2018).



Challenge 3: Causality is more complicated than we thought

- ▶ Getting a complete picture of causality is hard.
 - Existence: Does X cause Y ? If yes, what is the direction?
 - Importance: If yes, X have a non-trivial impact on Y ? How do we quantify the size of effect? How do we know an effect is big or small?
 - Mechanism: How and why does X affect Y ? **Causal mediation analysis** and (qualitative) **process tracing** are commonly used to unpack a causal mechanism.
- ▶ Tradeoff between **internal** and **external** validity: It is also hard to generalize from our findings.



Causal inference in action: Field and “natural” experiments

	Lab/Field Experiment	Natural Experiment
Comparing responses between treatment and control groups?	Yes	Yes
Does the treatment assignment occur at random?	Yes	More or less
Does the researcher control the introduction of the intervention (i.e., the treatment)?	Yes	No



Example: White et al (APSR 2015)

American Political Science Review

doi:10.1017/S0003055414000562

Vol. 109, No. 1 February 2015

© American Political Science Association 2014

What Do I Need to Vote? Bureaucratic Discretion and Discrimination by Local Election Officials

ARIEL R. WHITE *Harvard University*

NOAH L. NATHAN *Harvard University*

JULIE K. FALLER *Harvard University*

*D*o street-level bureaucrats discriminate in the services they provide to constituents? We use a field experiment to measure differential information provision about voting by local election administrators in the United States. We contact over 7,000 election officials in 48 states who are responsible for providing information to voters and implementing voter ID laws. We find that officials provide different information to potential voters of different putative ethnicities. Emails sent from Latino aliases are significantly less likely to receive any response from local election officials than non-Latino white aliases and receive responses of lower quality. This raises concerns about the effect of voter ID laws on access to the franchise and about bias in the provision of services by local bureaucrats more generally.

- ▶ Do local election officials in the US respond to voters of different ethnicities differently?
- ▶ 7,000 email requests were sent to local election offices; each request has a fictional voter from a uniquely "randomized" ethnicity (Latino v non-Latino).
- ▶ Emails sent from Latino aliases were less likely to receive a response or more likely to receive a response of lower quality.



Example: Cheng and Urpelainen (SCID 2019)

Studies in Comparative International Development (2019) 54:501–527
<https://doi.org/10.1007/s12116-019-09290-5>



Criminal Politicians and Socioeconomic Development: Evidence from Rural India

Chao-Yo Cheng¹ · Johannes Urpelainen²

- ▶ Does the election of criminal politicians undermine socioeconomic development in rural India?
- ▶ Analysis draws on fine-grained village-level (on local public goods and socioeconomic development) and constituency-level data (on candidate characteristics).
- ▶ Identification considers the victory of criminal politicians in close election "as-if" random.
- ▶ Criminal politicians may undermine household-level poverty alleviation while having no statistically discernible impact on infrastructure construction (e.g., paved roads).



How causal “revolution” changes social research

- ▶ Qualitative and multi-method researchers in demand
- ▶ Design-based "identification" precedes model-based inference
- ▶ Statistical learning: The rise of "causal" learning



How causal “revolution” changes social research

- ▶ Qualitative and multi-method researchers in demand
 - Qualitative information is important for research design and the search for proper setting for natural experiments (Dunning 2015).
 - It is possible to use qualitative methods/information to strengthen causal estimation (Glynn and Ichino 2015) and the study of mechanisms (Walder 2012).
- ▶ Design-based "identification" precedes model-based inference
- ▶ Statistical learning: The rise of "causal" learning



How causal “revolution” changes social research

- ▶ Qualitative and multi-method researchers in demand
- ▶ Design-based "identification" precedes model-based inference
 - The conventional approach is to include "all" possible predictors in multiple regression (i.e., the notorious "kitchen-sink" approach).
 - Design-based causal "identification" urges us to think hard about the distinction between the "cause" and pre-treatment/control variables.
- ▶ Statistical learning: The rise of "causal" learning



How causal “revolution” changes social research

- ▶ Qualitative and multi-method researchers in demand
- ▶ Design-based "identification" precedes model-based inference
- ▶ Statistical learning: The rise of "causal" learning
 - (Supervised) machine learning techniques may help with variable selection in multiple regression.
 - Statistical learning may also help with the analysis of "heterogeneous" causal effects.



What we offer at Birkbeck

- ▶ Advanced topics in quantitative social research (Cheng)
 - Generalized/non-linear multiple regression
 - Multilevel/hierarchical modeling
 - Basic tools for (observational) causal inference

- ▶ Experiments in social science (Maydom)
 - Lab and lab-in-the-field experiments
 - Survey experiments
 - Natural experiments

