Methods in Economic Research

Simon Naitram, PhD

Department of Economics, University of the West Indies

March 7, 2020

Section 1

Methodologies and Methods

The Journal of Economic Methodology says:

> *"The Journal distinguishes between methodology (which concerns the relationship between economics and broad questions about scientific knowledge) and methods (which involve particular techniques relevant to practitioners in a specific field of economics)"*

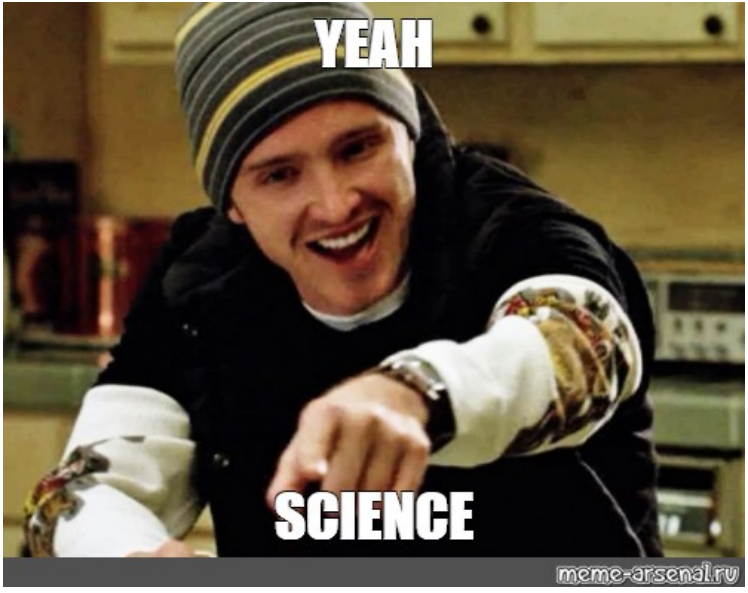
- Methodology is about broad philosophical questions about the relationship between economics and scientific knowledge.
- Method concerns more specific techniques used by, and primarily of interest to, practitioners in various fields of economics. (Hands, 2019)

Methodology versus Methods

- Methodology is the theoretical framework used to support the methods chosen.
- Methodology is the overarching strategy and rationale for your research.
- Methods are the specific tools and procedures you use to collect and analyse data.
- Methodology does not define specific methods, but helps you to identify the methods you should use.
- In shorter papers, we're mostly concerned with your methods. In longer studies (like a dissertation), we care about why you chose the approach you did.

The Economic Methodology

- Concerns the principles underlying economic reasoning.
- How do you choose what approach you're going to take to answer your question?
- Economists rarely think about this, and focus instead on getting on with research, focusing on specific methodological problems when they arise (Backhouse, 2016).
- There is no single general economic methodology, although there are a few principles that are shared by almost all fields: optimization, equilibrium, comparative statics.


YEAH
SCIENCE
meme-arsenal.ru

The Scientific Method

- The scientific method is an idealisation based on **empiricism**.
- This is in opposition to validating claims by appealing to:
  - revelation,
  - political or religious dogma,
  - tradition,
  - commonly held beliefs,
  - common sense,
  - currently held theories.
- **The scientific method is the recipe for how to do science!**

The Scientific Method

- Come up with question (e.g. by how much does an extra year of schooling increase lifetime earnings?)
- Formulate a hypothesis, usually based on pre-existing knowledge (e.g. theory)
  - A scientific hypothesis must be falsifiable
  - You must be able to identify an outcome that conflicts with your hypothesis
  - Otherwise this cannot be meaningfully tested.
- Predict/determine the logical consequences of your hypothesis.
  - Your predictions should be able to distinguish your hypothesis from likely alternative hypotheses.
  - Be concerned about observational equivalence.
- Test it, using some form of empirical analysis.
- Conduct analysis of your results.

Qualitative v Quantitative Research

- Qualitative approaches refer to research that is both descriptive and systematic (Mason, 2013).
  - Qualitative methods focus on groups, in-depth interviews, and reviews.
  - Qualitative methods do not rely on statistics or numbers and are less generalizable than quantitative methods.
  - Qualitative research approaches are subjective in nature.

- Quantitative approaches to research are based on formal, objective, and systematic processes in which data are numerically quantified (Mason, 2013).
  - Quantitative approaches are objective, deductive, and based on numeric quantification and generalization of results.
  - Quantitative research approaches are objective in nature.

Qualitative *and* Quantitative Research

- The two methods are symbiotic for the field, even though quantitative methods are more popular in economics
- The qualitative method allows you to explore ideas and experiences in depth, examining new areas, new people, new behaviours, and coming up with great theories and hypotheses.
- The quantitative method allows you to systematically test these hypotheses and generalise them to larger portions of the population.
- It is possible to combine these benefits in a single paper, using what we call a 'mixed methods approach'.

Mixed Methods Research

- You collect and analyse both qualitative and quantitative evidence in the same paper.
- Allows you to first examine diverse perspectives, developing testable hypotheses, then generalise them to a broader population and put a numerical point on the mechanism of interest.
- Or you can use quantitative data to uncover some pattern, then dig deeper to understand why it exists using deeper qualitative research.
- For example, consider a randomised control trial that collects quantitative data on women's choice of birth method and the resulting outcomes. Then collects qualitative data to better understand why they made these decisions.
- It adds to the richness, but also the complexity.

Abduction

- Abduction is the process of generating and choosing models, hypotheses and data analysed in response to surprising findings (Heckman and Singer, 2017).
- Abduction defines the research process as an iterative dynamic process, using multiple models and multiple sources of data in a back and forth method as learning and knowledge evolves.
- "The surprising fact, C, is observed. But if A were true, C would be a matter of course. Hence, there is reason to suspect that A is true." (Peirce, 1934)

Abduction—the detective approach

- Abduction is more a Sherlock Holmes approach.
- You use many different kinds of clues of varying trustworthiness, weigh them, put them together, and tell a plausible story of the ensemble.
- No need to privilege any specific kind of data or method, and the data does not need to be numerical.

Empirical Analysis

Economic empirical analysis typically has one of three goals:

- Advance knowledge by uncovering new facts or providing richer descriptions of old facts.
- Identify causal impacts of specific interventions.
- Understand the mechanisms producing outcomes with an eye toward interpretation and counterfactual policy evaluation.

Your study should do at least one of these.

Empirical Analysis

- We typically think of research as doing the latter two things: estimating causal effects or understanding mechanisms, typically through the use of regression methods.

- Note that there are lots of interesting studies that can simply **measure** economic quantities of interest.
  - The debate on the measurement of inequality is particularly heated in the US Saez and Zucman (2016).
  - Measuring factorless income or measuring missing factors of production Karabarbounis and Neiman (2019).

Methods for Collecting Data

- Surveys
- Focus groups
- Experiments
- Case studies
- Observational studies
- Secondary data analysis
- Online data collection and web scraping
- Meta-studies

Methods for collecting data

What matters when choosing a method for collecting data?

- Your question!
- Time and availability of resources.
- The level of detail you require.

Case Studies

- Case studies are almost never used in economics.
- There is, however, a generalisation that we use in economics: **comparative analysis**.
- Traditional aims of such analyses would be to examine the difference in outcomes between capitalist and socialist economies, and more recently, between different types of market economies.
- The underlying idea is that by comparing alternative economic systems, we can better understand what makes each of them work.
- Such analyses still need to be built on substantial amounts of data.

Surveys

- A survey is a list of questions aimed at extracting data from a group of people
- Surveys can be conducted by any medium you choose: telephone, online, in-person, mail, etc. (or even multiple modes)
- Survey questions can be open-ended or closed-ended (or partially-closed):
  - Open-ended questions allow respondents to respond based on their complete knowledge and understanding.
  - Closed-ended questions ask respondents to choose from a pre-defined list of answers.
  - Partially-closed

Surveys

Closed-ended questions can be asked with ordered or unordered categories:

- Scales (e.g. Likert scale, 1-5, best to worse, etc.),
- Yes or no questions
- Checklist multiple choice questions
- Rank order multiple choice questions

The benefit is that they limit respondent burden and are easy to tabulate, but the disadvantage is that they limit creativity or may lead respondents on.

Surveys

Open-ended questions typically try to elicit:

- Numerical responses
- List of items
- Description or elaboration

They're good for understanding the reasoning behind other answers, but are hard to analyse and puts greater burden on respondent.

Surveys

- The design of your survey is *the most important* part of a survey study.
- It is therefore important that you pre-test!
- Pre-testing is giving your survey to some test subjects and getting their feedback on what works and what doesn't.
- You then revise your survey (and maybe even pre-test again!)

Surveys

- Questions must be clear and unambiguous!
- Questions should not bias respondents!
- Differences in wording matters a lot!
- There are a billion do's and dont's; be simple, use common words, don't reinvent the wheel.
- If you plan on doing survey analysis, you will need to do a lot of reading to design your survey!

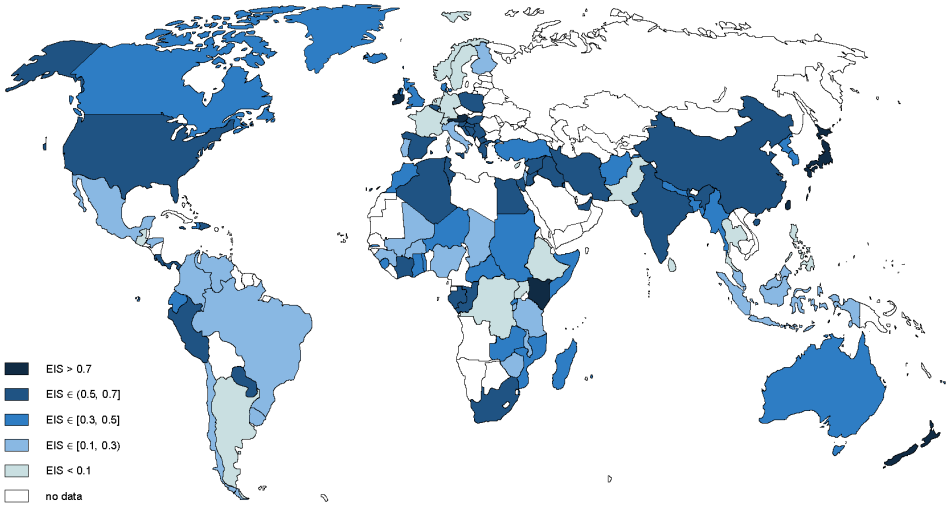
Interviews and questionnaires

- Interviews and questionnaires are the two broad types of ways you can administer surveys.
- The choice of how to administer your survey depends on the questions you want to ask.
- When there are more open-ended questions, prefer interviews.
- When you need a large number of respondents, prefer questionnaires.

Meta-analysis

- Meta-analysis is often forgotten, but has become very interesting and very rigorous recently.
- A meta-analysis is a statistical analysis that combines the results of multiple scientific studies.
- You use a meta-analysis when you're looking at a question that has been studied many times before, but where each estimate has some error attached to it.
- The simplest method is simply to find all these estimates and take a weighted average.
- Recent meta-analysis accounts for publication bias (vital!), the empirical design and specification of each study, and can even examine differences is responses across countries, time, or other characteristics.

The elasticity of inter-temporal substitution Havranek et al. (2015)



The Average Impact of Microcredit: Meager (2019)

- Combines evidence from seven randomized controlled trials in developing countries.
- Examines whether on average access to microfinance might foster entrepreneurship or potentially harm poor households by creating credit bubbles.
- Examines the average impact of microfinance on household business profit, expenditures, and revenues, and consumption.
- She finds that the average impact on these outcomes is quite small, only about a 5% increase on average.
- When accounting for uncertainty, these effects have a moderate to high probability of being zero.

Experiments: the gold standard

- The randomised controlled experiment or randomised controlled trial (RCT) is the gold standard for knowing what works.
- The main weapon in our search for truth is the **control group design**.
- We test a treatment by comparing a group who has received the treatment against a group that has not received the treatment.

Experiments: the gold standard

The components of an RCT (Greeno, 2002):

- Identify a specific treatment.
- Ensure there is a control condition applied to a control group.
- Assign participants to treated and control groups randomly.
- Carefully define and measure outcomes before and after treatment.

Quasi-experiments: the silver standard?

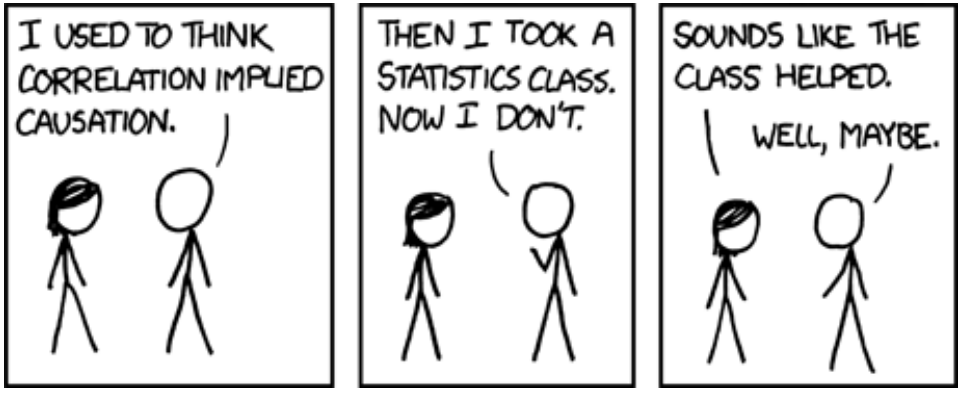
- Quasi-experimental designs use all the features of the RCT except the R!
- Participants are not randomly assigned, so the untreated group is now a 'comparison group' rather than a 'control group'.
- The drawback is that we no longer know whether the two groups are truly equivalent.
- The benefit is that experiments can be conducted in realistic settings.

Pre-test/post-test designs

- Studies like this do not compare different treatments.
- Instead, they simply compare the effect of a treatment on a participant before and after receiving the treatment.
- The drawback is that without a comparison group, the criticism arises that the same effect might have happened even if the participants didn't receive the treatment.
- You can also compare individuals who receive different treatments.

Section 2

Correlation and Causality


I USED TO THINK CORRELATION IMPLIED CAUSATION.
THEN I TOOK A STATISTICS CLASS. NOW I DON'T.
SOUNDS LIKE THE CLASS HELPED.
WELL, MAYBE.

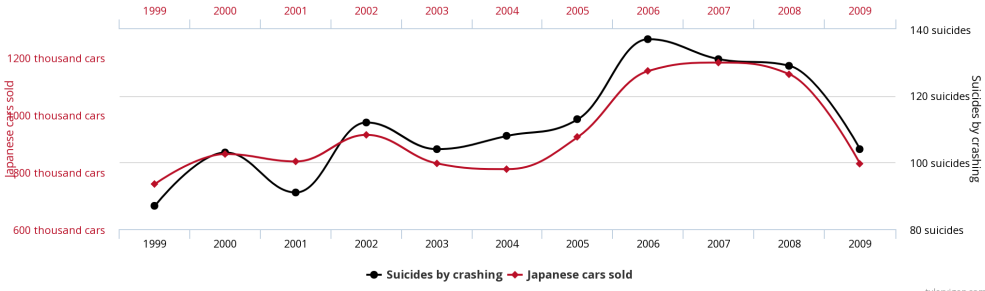
Validity

- Internal validity is the ability of the research to argue that the correlations you observe are causal.
  - Is not important in descriptive or exploratory studies.
  - It is very specific to your study and your research questions.
- External validity is the ability of the research to argue that your conclusions can be generalised to other people in other places at other times.
  - External validity depends on whether your sample has done a good job of representing the population.
  - You might even re-weight or stratify your sample so that the distribution of your sample looks like the distribution of the population.

Japanese passenger cars sold in the US

correlates with

Suicides by crashing of motor vehicle

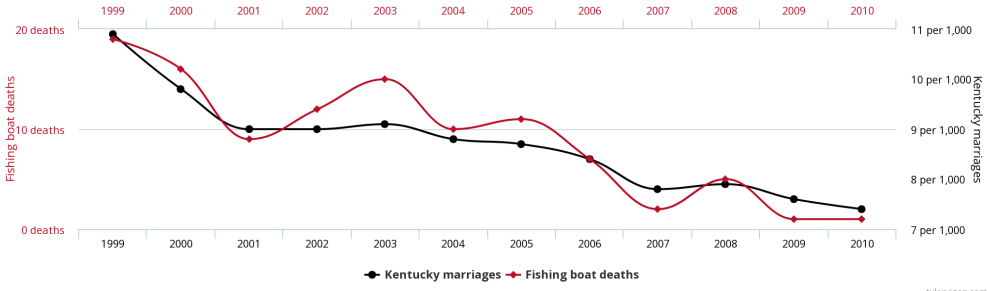


tylervigen.com

People who drowned after falling out of a fishing boat

correlates with

Marriage rate in Kentucky

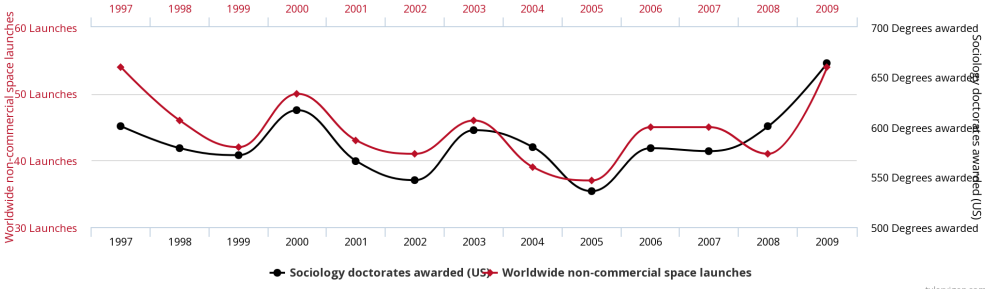


tylervigen.com

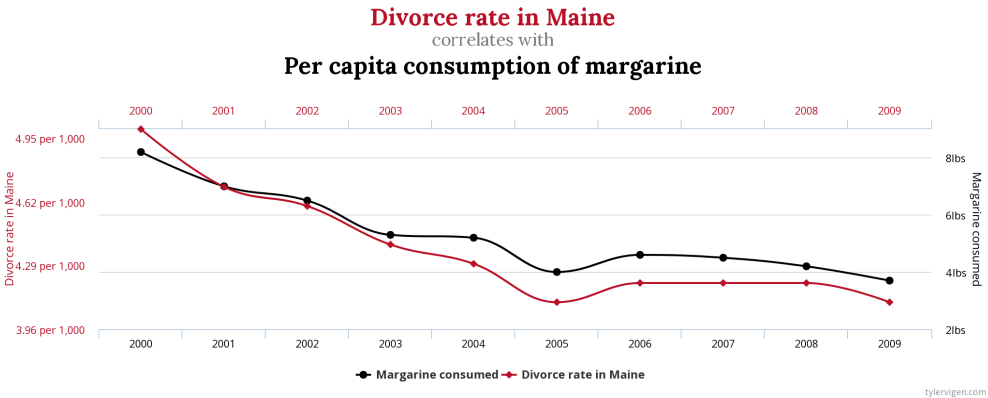
Worldwide non-commercial space launches

correlates with

Sociology doctorates awarded (US)



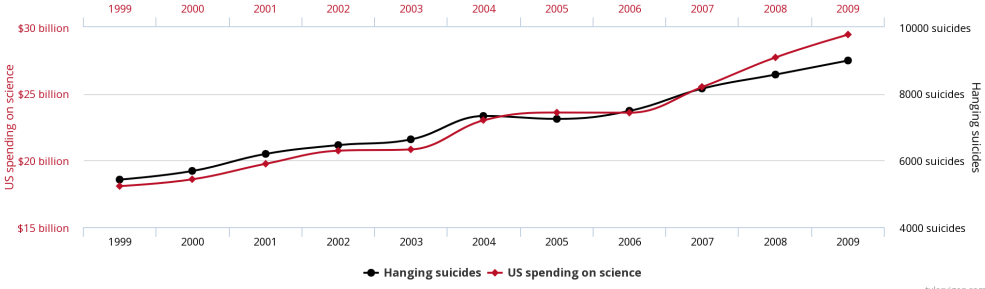
tylervigen.com


Divorce rate in Maine
correlates with
Per capita consumption of margarine
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
4.95 per 1,000
4.62 per 1,000
4.29 per 1,000
3.96 per 1,000
Divorce rate in Maine
8lbs
6lbs
4lbs
2lbs
Margarine consumed
Margarine consumed
Divorce rate in Maine
tylervigen.com

US spending on science, space, and technology

correlates with

Suicides by hanging, strangulation and suffocation



tylervigen.com

Potential explanations for observed correlations

- A and B are correlated because A causes B, or B causes A
- A and B are correlated because they are both caused by a third variable C
- A and B are correlated because A causes X, which, in turn, causes B
- A and B are correlated by pure chance, perhaps limited to the frame of observation alone

Section 3

Experiments in Economics

Randomized controlled trials

- We've always said you can't do experiments in economics, but after the mid 1990s, there was an explosion of RCT to inform economic policy.
- In 2019, Abhijit Banerjee, Esther Duflo and Michael Kremer won the Nobel prize in economics for their "experimental approach to alleviating global poverty".
- Experiments are no longer out of our reach, but are quite often costly.

The experimental approach

- More importantly for the reputation of economics as a science has been the adoption of the experimental approach.
- This has manifest itself in the 'empirical revolution' or the 'credibility revolution' (Angrist and Pischke, 2010).
- The primary mechanism driving the revolution has been a focus on the **quality of empirical research designs.**
- The aim has been to "take the con out of econometrics" (Leamer, 1983).

The experimental approach

- Quasi or natural experiments have been the main way in which economics has adopted the experimental method and improved the credibility of empirical economics.
- Natural experiments are random policies or events that give you treatment and control groups, and sometimes even random assignment.
- Even if you're not using these experimental-style methods, they help you to think of your specific problem in a clearer and more critical way.
- Each of the experimental approaches correspond to an empirical and/or statistical approach that we commonly see and you've learnt.

Natural experiments: examples

- Card (1990) used the Mariel boatlift—a mass immigration of Cubans to Miami in 1980—to estimate the effect of immigration on the Miami labour force. He finds that it increased the labour supply by 7%, but had virtually no effect on wages or unemployment rates of lower-skilled Miami residents.

- **?** used a televised draft lottery for the US army to estimate the effect of joining the army on future earnings. The lottery chose people born on a specific day to join the lottery. Those who weren't selected but born in the same year acted as the control. They found that among white men, drafted veterans went on to earn 15% less than their peers who had avoided the army.

Section 4

Quantitative Methods

Choice of Method

- The choice of method is determined by the data you have and the question you wish to answer.
- Unless you have explicitly done qualitative research, you will always have to use some quantitative methods (and even then, text analysis is a great quantitative way of analysing qualitative data).
- We'll discuss ways of approaching your analysis.

Causality

- Suppose we have the theoretical relationship

$$y = f(x).$$

- This is deterministic such that if we know x, we know y, using the function $f(.)$
- By saying y is a function of x, are we saying that this is a causal relationship? That x causes y?
- Notice that nothing prevents us from writing the equation as:

$$x = f^{-1}(y).$$

- Your default position should be to assume that any given correlation is simply correlation and not causal.
- In economic research, the **biggest problem** is always that your 'exogenous' regressor is not actually exogenous.
- There's always someone in the back of the room saying that your regressor is endogenous because of
  - reverse causality or simultaneity bias,
  - unobserved heterogeneity or omitted variables,
  - measurement error.

Causality

- Our typical regression looks like:

$$y = \alpha + \beta x + \varepsilon$$

- You can estimate this with:
  - Cross-sectional data
  - Time-series data
  - Panel data
  - Repeated cross-sectional data

- While the methods you choose will differ, you have to think about your method carefully in each case.

The Truth

- The truth is that the choice of econometric estimator is nowhere near as important as the basic stuff.
- You need to care about stuff like:
  - whether your sample has been chosen appropriately,
  - whether there is significant measurement error in your data,
  - have you defined/measured your variables correctly,
  - is your sample representative so that you can claim external validity.

Time Series

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

- Time series analysis is the study of dependence among the observations at different points in time.
- Aggregate economic variables are most likely to be represented as time-series data: inflation, GDP, wages, etc.
- With time series analysis, we are often concerned with the relationship between current and past values of the variables as well.
- Time series analysis is the language of pure dynamic economic theory.

Cross-Sectional

$$y_i = \alpha + \beta x_i + \varepsilon_i$$

- Cross-sectional analysis examines a population or sample of the population at a specific time.
- Surveys are a main source of cross-sectional data.
- With cross-sectional data, we are concerned with variation in outcomes across individuals.
- Cross-sectional analysis is the language of static economic theory.

Panel Data

$$y_{it} = \alpha + \beta x_{it} + \varepsilon_{it}$$

- Panel data is data on multiple individuals for multiple periods of time.
- Panel data is the best hope for establishing causality.
- Panel data allows us to consider many different empirical designs, and approaches the classical experimental methodology.
- However, it brings along baggage from both time-series and cross-sectional analysis.

Panel Data

- Allows you to examine variation both **over time** and **across individuals**.
- It allows us to eliminate the concern that individuals are different in permanent, unobservable ways.

$$y_{it} = \alpha_i + \beta x_{it} + \varepsilon_{it}$$

- It allows us to eliminate the concern that something funny was happening at this specific point in time to everyone.

$$y_{it} = \alpha_i + \delta_t + \beta x_{it} + \varepsilon_{it}$$

Panel Data

- You can get even fancier with it, eliminating group-specific time trends:

$$y_{ikt} = \alpha_i + \delta_k \cdot t + \beta x_{it} + \varepsilon_{it}$$

  which arises from a concern that different groups grow in different ways. For example, that tech firms grow faster on average than manufacturing firms.

- It even allows you eliminate the concern that there were funny things happening to specific groups at specific points in time:

$$y_{ikt} = \alpha_i + \delta_{kt} + \beta x_{it} + \varepsilon_{it}$$

  which are called group-specific common shocks.

Repeated Cross Sectional

- Repeated cross sectional data refers to data where a new random sample is collected at multiple time periods (eg. Population Census Data or Labour Force Survey).
- It means that it is either not on the same individuals, or you don't know whether the same individuals were surveyed both times.
- This is **not** panel data!
- But, we can sometimes aggregate this data into sub-populations to create a 'pseudo-panel'. This way we have changes in groups over time.
- Be careful, you may need to reweight the data or be cautious about changes in the survey methodology or sampling techniques over time.

Covariates

- Covariates can either be used as regressors of interest or simply as controls.
- When used as regressors of interest (ie. we care very much about the size, sign, and significance of the coefficient), you need to be able to make very clear and careful arguments that it is exogenous (statistically, not theoretically exogenous).
- Using covariates as controls allows you to control for confounding effects, or to make better comparisons by holding other characteristics constant (eg. controlling for income).

Section 5

Non-Regression Methods

Simple methods

- Sometimes we can get answers from methods that don't necessarily require regression analysis.
- In most cases, there is an underlying regression that helps you estimate the exact size of the effect and the size of the standard error (uncertainty of the estimate).
- These help to convince us that we're actually finding what we say we're finding, and are more compelling than complicated methods.

Difference in differences

- This is the simplest version of the experimental design.
- You need a treated group and a control group.
- You measure each group's average outcome *before* the treatment, and their average outcomes *after* the treatment.
- You calculate the change in each group's outcomes, after-before
- You calculate the *difference* in each group's change

Difference in differences: Chetty et al. (2009)

- Chetty et al. (2009) wanted to know whether people behaved differently when the listed price of a good included the sales tax versus if the sales tax was only added at the counter.
- In a single store for a few goods, they listed the price of the good inclusive of the sales tax.
- This is a nice experiment: you have a set of goods in the store that were treated, and a set of goods that weren't treated.
- The outcome of interest was the quantities sold before and after.

Difference in differences: Chetty et al. (2009)

TABLE 3— EFFECT OF POSTING TAX-INCLUSIVE PRICES: *DDD* ANALYSIS OF MEAN QUANTITY SOLD

Period	Control categories	Treated categories	Difference
Panel A. Treatment store			
Baseline (2005:1–2006:6)	26.48	25.17	−1.31
	(0.22)	(0.37)	(0.43)
	[5,510]	[754]	[6,264]
Experiment (2006:8–2006:10)	27.32	23.87	−3.45
	(0.87)	(1.02)	(0.64)
	[285]	[39]	[324]
Difference over time	0.84	−1.30	$DD_{TS} = -2.14$
	(0.75)	(0.92)	(0.68)
	[5,795]	[793]	[6,588]

Difference in differences in differences: Chetty et al. (2009)

- You can even increase how compelling this analysis is by adding another store.
- This time, the store isn't treated *at all*.
- You conduct the same difference-in-difference analysis on this store where no price tags were changes.
- Then you calculate the difference in the difference-in-difference estimates.
- This is a diff-in-diff-in-diff or a triple differences estimate.

Difference in differences: Chetty et al. (2009)

Panel B. Control stores			
Baseline (2005:1–2006:6)	30.57	27.94	-2.63
	(0.24)	(0.30)	(0.32)
	[11,020]	[1,508]	[12,528]
Experiment (2006:8–2006:10)	30.76	28.19	−2.57
	(0.72)	(1.06)	(1.09)
	[570]	[78]	[648]
Difference over time	0.19	0.25	$DD_{CS} = 0.06$
	(0.64)	(0.92)	(0.95)
	[11,590]	[1,586]	[13,176]
DDD Estimate			−2.20
			(0.59)
			[19,764]

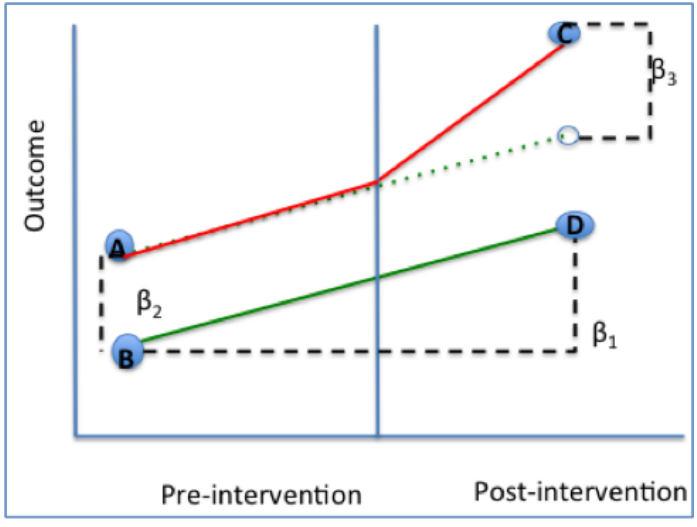
DD and DDD: Chetty et al. (2009)

- These estimates have a simple regression analogue.
- For example, the difference in difference estimate can be estimated by

$$y_{it} = \beta_0 + \beta_1 P_t + \beta_2 D_i + \beta_3 (D_i \times P_t) + \varepsilon_{it}$$

  where D_i is a dummy variable that is 1 if the product is treated, and P_t is a dummy variable that is 1 for the period(s) that the treatment was applied.
- The advantage with this approach is that you can add control variables and easily get standard errors.

DD in Pictures



DDD in Regression: Chetty et al. (2009)

TABLE 4—EFFECT OF POSTING TAX-INCLUSIVE PRICES: REGRESSION ESTIMATES

Dependent variable	Quantity per category (1)	Revenue per category ($) (2)	Log quantity per category (3)	Quantity per category (4)	Quantity (treat. categories only) (5)
Treatment	−2.20	−13.12	−0.101	−2.27	−1.55
	(0.60)	(4.89)	(0.03)	(0.60)	(0.35)
Average price	−3.15	−3.24		−3.04	−15.06
	(0.26)	(1.74)		(0.25)	(3.55)
Average price squared	0.05	0.06		0.05	1.24
	(0.00)	(0.03)		(0.00)	(0.34)
Log average price			−1.59		
			(0.11)		
Before treatment				−0.21	
				(1.07)	
After treatment				0.20	
				(0.78)	
Category, store, week FEs	x	x	x	x	x
Sample size	19,764	19,764	18,827	21,060	2,379

Regression Discontinuity Design (RDD)

- Government policies often create sharp cut-off points:
  - Change in retirement age for those born after April 1, 1950; but not for those born just before.
  - Students just above 3.6 GPA get first-class honours but those just below don't.
  - People living on the same street might be grouped into different school districts because the border falls between them.

- It is argued that individuals cannot control their assignment into either group when they're very near this cut-off, so their assignment to being treated or not is effectively random.

- That is, it looks just like a randomised experiment!

Regression Discontinuity Design (RDD)

- With RDD, the analysis is very simple.
- Draw two lines of best fit in the x-y plane: one for those not treated and one for those treated.
- You can play with the types of best fit lines that are drawn: linear/nonlinear, same slope but different intercept, etc.
- The size of the treatment effect is the distance between where the two lines intercept your cut-off point.

RDD Examples

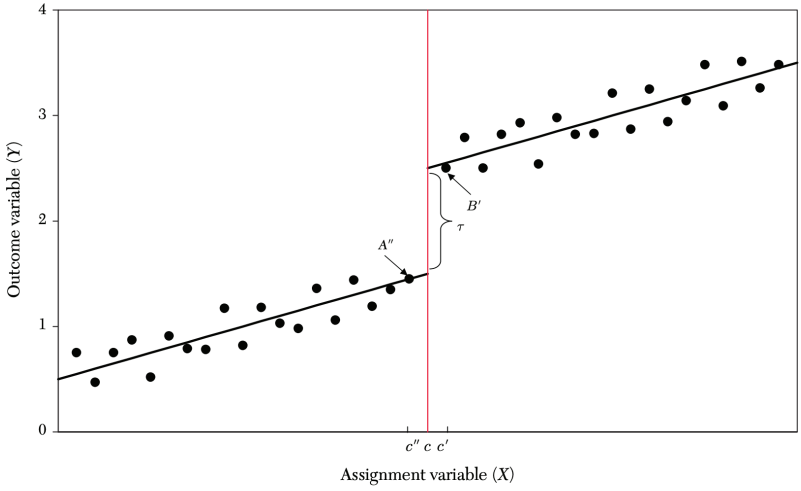
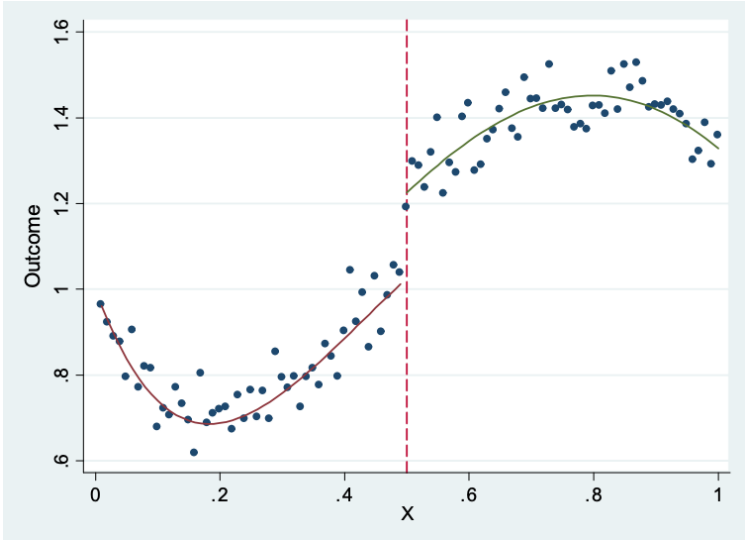


Figure 1. Simple Linear RD Setup

RDD Examples



Regression Discontinuity Design (RDD)

- And as before, there's a really simple way to estimate the size of this difference using a regression:

$$y_i = \alpha + \tau D_i + \beta X_i + \varepsilon_i$$

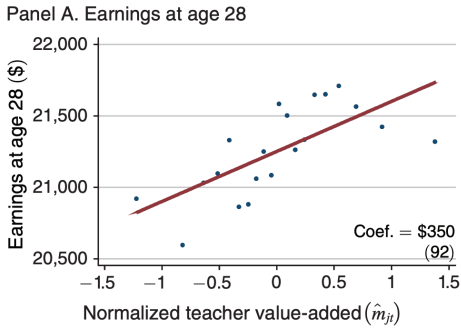
where D is 1 if you're above the cutoff and 0 if you're below the cutoff. Then τ is the effect of the treatment at the cutoff.

Binned Scatterplots

- The binned scatterplot is a nice way to plot the relationship between two variables when you have lots of (possibly noisy) data.
- Effectively all you're doing is grouping the x variable into bins (groups based on the values of x).
- Then take the average y value of each group, and plot the average of each x group against group's average y.
- In Stata, the command binscatter can control for covariates before plotting the relationship between x and y.

Binned Scatterplot

- What is the effect of teacher quality on student's future earnings?
- Chetty et al. (2014) use the impact of teachers on students' test scores to measure teacher quality.
- This graph visually represents a multivariate regression on 650,965 observations, finding that a 1 unit increase in normalized teacher value-added increases future earnings by $350 per year.
- Note that this actually includes a wide range of controls.



Event studies

- One really popular method in finance and financial economics is the event study.
- Event studies are a simple way to test the impacts of public policy on market outcomes by summing the unexpected returns as new information meets the market.
- They're really useful in examining the responses to political events as well.
- The main problem is that quite often these events aren't entirely unexpected, and we have to adjust for the predicted probability of the event happening.
- The analysis is usually concentrated in really short time periods (days, hours, minutes).

Event studies: example

- What was the effect of Obamacare on expected returns for different healthcare firms?
- The election of a Republican to the senate in a special election in Massachusetts to replace a Democrat substantially reduced the likelihood of the bill passing.
- This win was unexpected, and his election had a *positive* effect on healthcare stocks.
- A typical dollar invested in the health care sector realized a 2.1 percent Cumulative Abnormal Return (CAR) between January 14, 2010 and January 20, 2010.

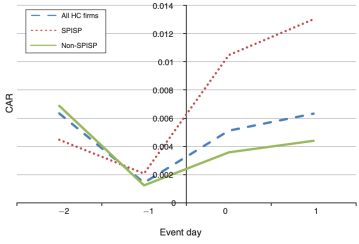


Figure 2. Cumulative Abnormal Return (Equally Weighted) by Event Day for All Health Care Firms, Health Care Firms in the SPISP, and Health Care Firms Not in the SPISP

Section 6

Tips and Tricks of the Trade

Tips and Tricks of the Trade

- Robustness checks are important! Re-run your model as many times as you need with various specifications to *convince yourself*. Present as many of these as possible.
- Ask questions that are interesting regardless of the result you get. This ways you won't want to fight the data.
- Be careful with what covariates you use. Sometimes they can spuriously invent correlations!

Tips and Tricks of the Trade

- Economists typically estimate variables in natural logs, because the coeffcients give us elasticities.

$$\beta = \frac{\partial \ln(y)}{\partial \ln(x)} \approx \frac{\partial y}{\partial x}\frac{y}{x}$$

- Be careful with natural logs however, since any zero or negative values will be dropped. In this case, you should use the inverse hyperbolic sine (arcsinh):

$$\text{arcsinh}(x) = \ln\left(x + \sqrt{x^2 + 1}\right)$$

  which keeps zero and negative values. Bellemare and Wichman (2019) shows how to back out the point elasticity. It's very simple.

- A semi-elasticity is when the independent variable is not logged. It is interpreted as "a 1-unit change in the independent variable is correlated with a β percent change in the dependent variable."

Tips and Tricks of the Trade

- Your R-squares isn't really important. However, if you're making a big deal about the effect of x on y, and your x variable explains very very little (e.g. 0.01) variation in y, then maybe it isn't quite the magic policy instrument that you are looking for. It means there are likely other things that would do a better job in changing y (if this is what policymakers are interested in).

- When you have serious questions about whether you should do thing A or thing B, you should probably do both. For example, should I cluster my standard errors or not? Do both. Should I estimate this in levels or in first-differences? Do both. There's no single answer, and in economics, skin your cat in all the ways you can find.

Tips and Tricks of the Trade

- Plot your data. A lot. Slice it and dice it in every which way possible. Look at it from every angle. You'll learn things about your data, and you'll uncover mistakes, outliers, and data inconsistencies like this. In addition, presenting plots helps your reader understand what you're doing. If a picture is worth a thousand words, then you can keep within the word limit with a few nice images.

- When cleaning and transforming your data, make sure you document **everything** you do. It's best to do it by writing code (using Stata or R or something), and if you use Excel to do anything with your data, make sure you write down every single step you do. You're writing a chemistry lab report which another chemist could use to replicate your work.

Tips and Tricks of the Trade

- Statistical significance isn't economic significance. Just because an effect is statistically different from zero doesn't mean that it is big in real-world terms. A 1 percent increase in car prices might lead to a 0.001 percent decrease in the quantity of cars bought. This effect isn't economically significant. Discuss both statistical and economic significance of your estimates. Be honest and up-front about it.

- It is commonly viewed as very important that you cluster your standard errors when you have panel data in particular.

- Consider whether you need to weight your estimates (eg. if you have a panel of countries and some are huge and some are small).

Section 7

Further sources

- https://guides.lib.vt.edu/researchmethods/design-method
- https://en.wikipedia.org/wiki/Scientific_method
- https://evidencebasedliving.human.cornell.edu/2011/04/23/randomized-controlled-designs-the-gold-standard-for-knowing-what-works
- https://socialresearchmethods.net/kb/constructing-survey/
- http://meta-analysis.cz/
- http://economicspsychologypolicy.blogspot.com/2015/06/list-of-19-natural-experiments.html

- http://nickchk.com/causalgraphs.html
- https://michaelstepner.com/binscatter/
- https://sites.google.com/site/andrewjohnstoneconomics/my-advice/event-studies
- https://data.nber.org/reporter/fall02/newEconomics.html

Section 8

References

Angrist, J. D. and J.-S. Pischke (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of economic perspectives 24*(2), 3–30.

Backhouse, R. E. (2016). Methodology of Economics BT - The New Palgrave Dictionary of Economics. pp. 1–7. London: Palgrave Macmillan UK.

Bellemare, M. F. and C. J. Wichman (2019). Elasticities and the Inverse Hyperbolic Sine Transformation. *Oxford Bulletin of Economics and Statistics forthcomin*.

Card, D. (1990). The impact of the Mariel boatlift on the Miami labor market. *ILR Review 43*(2), 245–257.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American economic review 104*(9), 2633–2679.

Chetty, R., A. Looney, and K. Kroft (2009). Salience and taxation: Theory and evidence. *American economic review 99*(4), 1145–1177.

Greeno, C. G. (2002). Major alternatives to the classic experimental design. *Family process 41*(4), 733.

Hands, D. W. (2019). Economic Methodology in the Twenty-First Century (So Far): Some Post-Reflection Reflections. *Available at SSRN 3464759*.

Havranek, T., R. Horvath, Z. Irsova, and M. Rusnak (2015). Cross-country heterogeneity in intertemporal substitution. *Journal of International Economics 96*(1), 100–118.

Heckman, J. J. and B. Singer (2017). Abducting Economics. *American Economic Review 107*(5), 298–302.

Karabarbounis, L. and B. Neiman (2019). Accounting for factorless income. *NBER Macroeconomics Annual 33*(1), 167–228.

Leamer, E. E. (1983). Let's take the con out of econometrics. *The American Economic Review 73*(1), 31–43.

Mason, S. A. (2013). Qualitative Versus Quantitative Approaches BT - Encyclopedia of Autism Spectrum Disorders. pp. 2485–2486. New York, NY: Springer New York.

Meager, R. (2019). Understanding the average impact of microcredit expansions: A Bayesian hierarchical analysis of seven randomized experiments. *American Economic Journal: Applied Economics 11*(1), 57–91.

Peirce, C. S. (1934). Belief and judgment. *Collected papers 5*, 376–387.

Saez, E. and G. Zucman (2016). Wealth inequality in the United States since 1913: Evidence from capitalized income tax data. *The Quarterly Journal of Economics 131*(2), 519–578.