RESEARCH · RESEARCH DESIGN

Natural Experiments in Business and Social Research

What natural experiments are, where they come from (lotteries, rules, borders, timing and shocks), famous examples from public health, economics and business, how to judge whether assignment really was as good as random, and how organisations can find natural experiments in their own data.

ModQu research article · 1,559 words · 15 references (APA 7th edition)

Some of the most convincing evidence about cause and effect comes from experiments that no researcher designed. A lottery decides who is called up for military service. A policy changes on one side of a border but not the other. A new technology reaches some towns months before others. In each case, events outside the researcher's control have divided people or places into groups in a way that is unrelated, or nearly unrelated, to the outcomes being studied. These situations are called **natural experiments**. This article explains what makes a natural experiment, where they come from, how to judge them, and how businesses can use the same logic.

What makes an experiment "natural"

In a randomised experiment, the researcher uses chance to decide who receives a treatment, which makes the treated and untreated groups alike in every respect except the treatment. In a natural experiment, the assignment is made by nature, by an institution or by circumstance, and the researcher argues that it was **as good as random**: unrelated to the characteristics of the units that might otherwise affect the outcome (Dunning, 2012).

Meyer (1995) called the best of these situations ones in which an exogenous event, such as a policy change, shifts treatment for some units and not others. This distinguishes a natural experiment from an ordinary observational comparison. Comparing people who chose to take a training course with people who did not is not a natural experiment, because the choice reflects motivation, ability and circumstances that also affect earnings. Comparing people who were admitted to an oversubscribed course by lottery with those who were not is a natural experiment, because the lottery, not the person, decided.

Dunning (2012) emphasises that the strength of a natural experiment lies in the design, not in statistical modelling. When assignment is as good as random, a simple comparison of means between groups can be credible. When it is not, no amount of statistical adjustment fully rescues the study.

A classic example

The best-known early natural experiment comes from John Snow's investigation of cholera in London. In the 1850s, two water companies supplied houses in the same districts of south London, with pipes from both companies running down the same streets. One company drew water from a stretch of the Thames contaminated with sewage; the other had moved its intake upstream. Rich and poor houses were served by both. Snow (1855) found that cholera deaths were far higher among households supplied by the contaminated company. Because the two groups were otherwise similar, the comparison provided strong evidence that cholera was spread through water, decades before the bacterium was identified.

Where natural experiments come from

Natural experiments arise from several kinds of event. Each produces a different kind of comparison and rests on a different assumption.

Where natural experiments come from

Each source creates a comparison, and each rests on a different assumption

Source	Example	The comparison holds if...
Lottery	Draft lottery decides who is called for military service	The lottery was run fairly and followed
Rule or threshold	School-leaving age means birth date affects years of schooling	People cannot sort themselves around the rule
Border	Minimum wage rises in one state but not its neighbour	Both sides would have followed the same trend
Timing or rollout	Mobile coverage reaches coastal regions at different times	Rollout order is unrelated to local conditions
Shock	A sudden wave of migration enlarges one city's workforce	Comparison cities are similar and unaffected

Examples from Angrist (1990); Angrist & Krueger (1991); Card & Krueger (1994); Jensen (2007); Card (1990) ModQu.com

Figure 1. Five common sources of natural experiments, with an example of each and the assumption on which the comparison depends. Examples from Angrist (1990), Angrist and Krueger (1991), Card and Krueger (1994), Jensen (2007) and Card (1990); table by ModQu.

Lotteries

When an institution allocates something by lottery, it creates a true random assignment that a researcher can exploit later. Angrist (1990) studied the United States draft lotteries of the Vietnam War era, in which birth dates were drawn at random to decide the order in which men were called for service. Because birth dates are unrelated to earning potential, comparing men with high and low lottery numbers showed the effect of military service on later earnings, which turned out to be negative for white veterans.

Rules and thresholds

Many institutions apply rules that create sharp differences between otherwise similar people. Angrist and Krueger (1991) used the fact that compulsory schooling laws allowed students to leave school at a fixed age. Because students born early in the year reached that age after fewer years in school, quarter of birth slightly affected how much education people received, and this could be used to estimate the effect of schooling on earnings. Threshold rules of this kind are the basis of regression discontinuity designs.

Geography and borders

Policies often stop at a border, while the people and economies on either side remain similar. Card and Krueger (1994) compared fast-food restaurants in New Jersey, which raised its minimum wage, with those in neighbouring Pennsylvania, which did not. The comparison relies on the two areas following similar trends in the absence of the policy.

Timing and rollouts

When a technology, service or policy reaches different places at different times for reasons unrelated to local conditions, the staggered rollout creates natural comparison groups. Jensen (2007) studied the arrival of mobile phone coverage along the coast of Kerala, India, in the late 1990s. As coverage reached each region, fishermen could call markets while still at sea and sell where prices were best. Price differences between markets fell sharply and waste declined, region by region, as coverage arrived.

Shocks

Sudden, unexpected events can also create comparisons. Card (1990) studied the Mariel boatlift of 1980, when about 125,000 Cuban migrants arrived in Miami within a few months, expanding the city's labour force by around 7%. He found little effect on existing workers' wages or unemployment, compared with similar cities. Costello et al. (2003) studied a rural community in North Carolina in which a casino opened on a reservation and began paying a share of profits to every Eastern Cherokee family. Because the payments did not depend on family behaviour, they showed how moving out of poverty affected children's mental health.

Natural experiments in business

Businesses generate natural experiments constantly, often without noticing. A system outage disables a recommendation engine for half a day. A supplier problem leaves some stores without a product for a month. A competitor enters some markets and not others. A tax change alters prices in one jurisdiction. Einav et al. (2014) used listing data from eBay to study how sales taxes affect online shopping: because buyers in the seller's own state paid sales tax while other buyers did not, the same item listing faced different effective prices for different buyers.

Organisations can look for such events deliberately, asking whether anything changed suddenly for some customers, stores or regions and not others, and whether the reason for the change was unrelated to the outcome of interest. A planned rollout that proceeds by alphabetical order of store code, or by the order in which installation teams become available, may be close to random. A rollout that starts with the best-performing stores is not.

Judging a natural experiment

The central question is always whether assignment really was as good as random. Sekhon and Titiunik (2012) warned that many studies described as natural experiments are neither natural nor experiments: the assignment may have been influenced by the units themselves, or the treated and untreated groups may not be comparable even if assignment was random. Several checks help.

- **Understand the assignment process.** Find out exactly how and why units ended up in each group. Qualitative knowledge of the institution is often more informative than statistical tests (Dunning, 2012).
- **Check balance.** Compare the groups on characteristics measured before the event. If assignment was as good as random, they should be similar.
- **Check that the event changed only the treatment.** A border policy may coincide with other differences between jurisdictions; a shock may affect outcomes through several channels at once.
- **Be clear about whom the result applies to.** When a natural experiment affects the treatment only for some units, such as men whose service was determined by the lottery, the estimate applies to those units. Imbens and Angrist (1994) showed that such estimates describe the effect for the group whose behaviour the event changed, which may differ from the effect for everyone.

Rosenzweig and Wolpin (2000) added that natural experiments still require assumptions, often unstated, about how people respond. A natural experiment can establish an effect convincingly without explaining why it occurred.

Strengths and limits

Natural experiments combine two advantages that are rarely found together: credible causal comparison and real behaviour in real settings, often at a scale no researcher could fund. Craig et al. (2017) describe them as an essential tool in public health, where many important interventions, such as taxes, bans and planning decisions, cannot be randomised.

Their limits follow from the fact that the researcher cannot choose the treatment. Researchers must study the events that happen to occur, which may not match the question a decision-maker most needs answered. The groups affected may be unusual, the samples may be small, and the event may be unique. Angrist and Pischke (2009) argue that a good natural experiment is often worth more than a large but poorly identified dataset, but it is still one study in one setting, and its findings gain strength when similar events elsewhere produce similar results.

Conclusion

Natural experiments allow researchers to study causes that cannot be manipulated, by finding situations in which the world has done the randomising. Their value depends entirely on how close the assignment came to random, which must be argued from knowledge of the specific case and checked with data. For organisations, the practical lesson is to notice the sudden, uneven changes that happen in the course of business and ask whether they offer a comparison that a planned study could not.

References

Angrist, J. D. (1990). Lifetime earnings and the Vietnam era draft lottery: Evidence from Social Security administrative records. *American Economic Review, 80*(3), 313–336.

Angrist, J. D., & Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *Quarterly Journal of Economics, 106*(4), 979–1014. https://doi.org/10.2307/2937954

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Card, D. (1990). The impact of the Mariel boatlift on the Miami labor market. *Industrial and Labor Relations Review, 43*(2), 245–257. https://doi.org/10.1177/001979399004300205

Card, D., & Krueger, A. B. (1994). Minimum wages and employment: A case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review, 84*(4), 772–793.

Costello, E. J., Compton, S. N., Keeler, G., & Angold, A. (2003). Relationships between poverty and psychopathology: A natural experiment. *JAMA, 290*(15), 2023–2029. https://doi.org/10.1001/jama.290.15.2023

Craig, P., Katikireddi, S. V., Leyland, A., & Popham, F. (2017). Natural experiments: An overview of methods, approaches, and contributions to public health intervention research. *Annual Review of Public Health, 38*, 39–56. https://doi.org/10.1146/annurev-publhealth-031816-044327

Dunning, T. (2012). *Natural experiments in the social sciences: A design-based approach*. Cambridge University Press. https://doi.org/10.1017/CBO9781139084444

Einav, L., Knoepfle, D., Levin, J., & Sundaresan, N. (2014). Sales taxes and internet commerce. *American Economic Review, 104*(1), 1–26. https://doi.org/10.1257/aer.104.1.1

Imbens, G. W., & Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica, 62*(2), 467–475. https://doi.org/10.2307/2951620

Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *Quarterly Journal of Economics*, *122*(3), 879–924. https://doi.org/10.1162/qjec.122.3.879

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics, 13*(2), 151–161. https://doi.org/10.1080/07350015.1995.10524589

Rosenzweig, M. R., & Wolpin, K. I. (2000). Natural "natural experiments" in economics. *Journal of Economic Literature*, *38*(4), 827–874. https://doi.org/10.1257/jel.38.4.827

Sekhon, J. S., & Titiunik, R. (2012). When natural experiments are neither natural nor experiments. *American Political Science Review, 106*(1), 35–57. https://doi.org/10.1017/S0003055411000542

Snow, J. (1855). *On the mode of communication of cholera* (2nd ed.). John Churchill.