Do We Need 'Natural Experiments'?

Framework for natural experimental evaluations: Launch event

Sam Harper

2025-05-23

McGill | Department of Epidemiology, Biostatistics and Occupational Health

Strengths of the Framework

- Inclusive development with stakeholders
- Emphasis on theory, planning, context, transparency
- Importance of understanding treatment assignment mechanism
- Integration of mixed methods

Defining “Natural Experiments”

2012:

> The common thread in most definitions is that exposure to the event or intervention of interest has not been manipulated by the researcher.

2025:

> We use the term to refer to events or processes outside the control of a researcher that divide a population into exposed or unexposed groups, or groups with differing degrees of exposure.

See Craig et al. (2012); Craig et al. (2025)

Natural Experiments are Observational Studies

Paul Rosenbaum in *Observational Studies*:

> An observational study is an empiric investigation of the effects caused by a treatment, policy, or intervention in which it is not possible to assign subjects at random to treatment or control, as would be done in a controlled experiment.

Same, right? Is this a problem?

See Rosenbaum (2002)

Consumers of evidence likely make a distinction.

For example, GRADE guidance:

> In public health, where randomized studies are less common and often infeasible in comparison with other areas of health, some types of NRS may provide greater certainty than others when investigating the health effects of policy or social interventions. For example, natural experiment studies may address selection bias and confounding through designs such as ITS or regression discontinuity, which may support stronger causal inference than other observational designs such as cohort and case-control studies [45].

See Hilton Boon et al. (2021) and note that reference [45] is to Craig et al. (2017).

Potential for ‘halo effects’?

Preventive Medicine 116 (2018) 19–26

Contents lists available at ScienceDirect

Preventive Medicine

journal homepage: www.elsevier.com/locate/ypmed

Organised population-based programmes and change in socioeconomic inequalities in mammography screening: A 1992–2012 nationwide quasi-experimental study

Stéphane Cullati[a,b,c,d,*,1], Martina von Arx[d,1], Delphine S. Courvoisier[c], José Luis Sandoval[a,c], Orly Manor[e], Claudine Burton-Jeangros[b,d], Christine Bouchardy[f], Idris Guessous[a,g]

[a] Unit of Population Epidemiology, Department of Community Medicine, Primary Care and Emergency Medicine, Geneva University Hospitals, Switzerland
[b] Swiss NCCR “LIVES - Overcoming Vulnerability: Life Course Perspectives”, University of Geneva, Switzerland
[c] Department of General Internal Medicine, Rehabilitation and Geriatrics, University of Geneva, Switzerland
[d] Institute of Demography and Socioeconomics, University of Geneva, Switzerland
[e] School of Public Health and Community Medicine, Hebrew University-Hadassah, Israel
[f] Geneva Cancer Registry, Global Health Institute, University of Geneva, Switzerland
[g] Department of Ambulatory Care and Community Medicine, University of Lausanne, Switzerland

What makes this “quasi-experimental”?

> ...some Swiss regions do have organised breast cancer programmes, while others still rely on opportunistic screening.

> This ecological quasi-experimental context allows analysing the evolution of socioeconomic inequalities in mammography screening over time in the different regions.

- No discussion of treatment assignment mechanism
- No discussion of potential biases of the treatment effect

See Cullati et al. (2018)

Cluster-level treatment, but individual 'controls'

2.5. Control variables

Due to their potential associations with mammography screening (Cabeza et al., 2007; Carrasco-Garrido et al., 2014; Chamot and Perneger, 2003; Documet et al., 2015; Eisinger et al., 2015; Kang et al., 2014; Martín-López et al., 2013; Maruthur et al., 2009), the following variables were used as covariates: age (50–64, 65–70), nationality (Swiss, not Swiss), region of residence according to language (German, French, Italian), marital status (single, married, widow, divorced or separated), number of people living in the household (1, 2 or more), number of children 14 years old or younger living in the household (yes, no), type of insurance coverage for hospital stay (standard, half-private and fully private), number of close relations providing emotional support (many people, one person, no), self-rated health (very bad, bad, so-so, good, very good), body mass index (underweight, normal weight, overweight, obese), physical symptoms (no or a few, some, important), currently smoking (yes, no), visit to a general practitioner or family doctor in the last 12 months (yes, no).

Cullati et al. (2018)

Weak design, but causal conclusions?

> our results showed that organised screening programmes modified income inequalities by reducing differences towards the null.

> Despite the lack of statistical significance of APRs in both groups, it suggested that organised programmes reduced the APR of income towards the null.

> these programmes reversed the gap in mammography screening uptake between employed and not employed women (the latest exceeded the prevalence of employed women by 2012) and attenuated educational and income-related inequalities.

Cullati et al. (2018)

Eur J Health Econ (2017) 18:649–665
DOI 10.1007/s10198-016-0845-7

ORIGINAL PAPER

The effects of organized screening programs on the demand for mammography in Switzerland

Mark Pletscher[1,2]

✉ Mark Pletscher
markpletscher@gmx.ch

[1] Institute of Economic Research, University of Neuchâtel, Rue A.-L. Breguet 2, 2000 Neuchâtel, Switzerland

[2] Winterthur Institute of Health Economics, Zurich University of Applied Sciences, Gertrudstrasse 15, 8400 Winterthur, Switzerland

What makes this “quasi-experimental”?

- Exact same data, and clear objective:

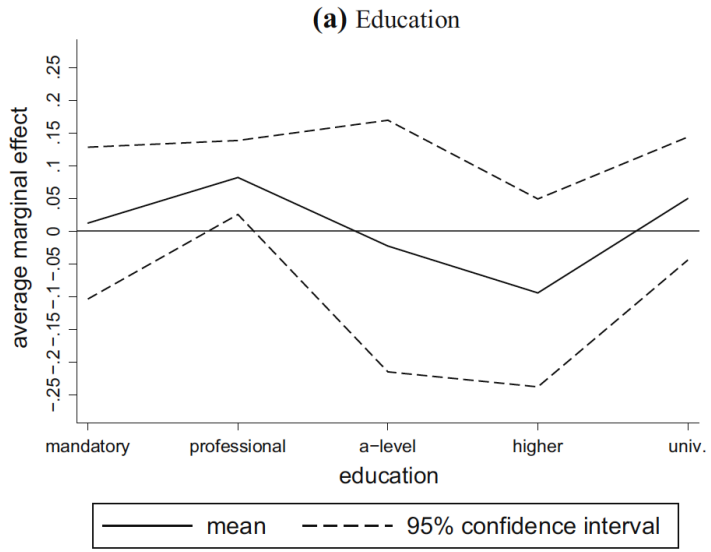
> to estimate the effect of organized mammography screening programs on screening initiation in screening cantons.

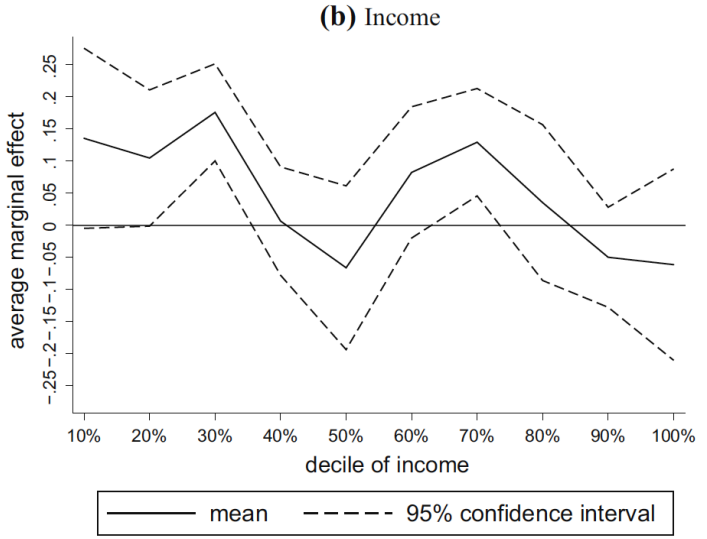
- Concerns about identification:
  - Include region and time fixed effects (i.e., DiD)
  - Functional form of model
- Evaluating alternative explanations by design:
  - Placebo tests on pre-intervention trends
  - Triple differences model

See Pletscher (2016)

Ambiguous evidence, sensible conclusions

> we do not find clear income- or education-related gradients in screening initiation. Although we find...stronger effects of the organized programs among women with lower incomes, these gradients are rather moderate.





Pletscher (2016)

Current thoughts (with an open mind)

A definition that fails to distinguish between stronger and weaker observational designs creates confusion.

Two suggestions:

1. Use a more stringent definition

For example, designs that ‘select on unobservables’

2. Drop the label entirely.

Focus on question, design, assumptions, and inference.


Questions?
sam.harper@mcgill.ca
samharper.org

References

Craig, P., Campbell, M., Deidda, M., Dundas, R., Green, J., Katikireddi, S.V., Lewsey, J., Ogilvie, D., Vocht, F. de, White, M., 2025. BMJ e080505.

Craig, P., Cooper, C., Gunnell, D., Haw, S., Lawson, K., Macintyre, S., Ogilvie, D., Petticrew, M., Reeves, B., Sutton, M., Thompson, S., 2012. Journal of Epidemiology and Community Health 66, 1182–1186.

Craig, P., Katikireddi, S.V., Leyland, A., Popham, F., 2017. Annual Review of Public Health 38, 39–56.

Cullati, S., Arx, M. von, Courvoisier, D.S., Sandoval, J.L., Manor, O., Burton-Jeangros, C., Bouchardy, C., Guessous, I., 2018. Preventive Medicine 116, 19–26.

Hilton Boon, M., Thomson, H., Shaw, B., Akl, E.A., Lhachimi, S.K., López-Alcalde, J., Klugar, M., Choi, L., Saz-Parkinson, Z., Mustafa, R.A., Langendam, M.W., Crane, O., Morgan, R.L., Rehfuess, E., Johnston, B.C., Chong, L.Y., Guyatt, G.H., Schünemann, H.J., Katikireddi, S.V., 2021. Journal of Clinical Epidemiology 135, 42–53.

Pletscher, M., 2016. The European Journal of Health Economics 18, 649–665.

Rosenbaum, P.R., 2002. Observational studies. Springer New York.