

From the research bench

What is causal inference?

Causal inference estimates what a change did by comparing outcomes that are never all observed. The design and its assumptions carry the answer.

Author	Published	Reading time	Words
Sinan Isoglu, MBA (Quantic) Commercial growth leader, lecturer and doctoral researcher	18 October 2026	17 min	3,767

Management summary

Causal inference is the estimation of what a treatment changed: the difference between the outcome a unit had and the outcome it would have had under the alternative. This concept guide sets out the potential outcomes model (Rubin; Rosenbaum and Rubin), explains why a raw comparison of treated and untreated groups is biased, and compares five routes to an estimate: randomized experiments, adjustment for observed covariates, natural experiments, fixed effects and difference-in-differences, and instrumental variables, with what each must assume. It covers estimands, causal diagrams (Pearl) and a one-page record for writing a causal claim down, and it takes its limits from Gangl's review and from those Pearl states for his own method.

Contents	§ 1	What is causal inference?	2
	§ 2	Why is correlation not causation?	3
	§ 3	Why do randomized experiments come first?	4
	§ 4	How can observational data support a causal claim?	5
	§ 5	What do natural experiments and repeated observations add?	6
	Figure 1	Five routes to a causal estimate, and what each must assume	7
	§ 6	Which assumptions travel with a causal claim?	10
	§ 7	Which effect is being estimated?	11
	§ 8	How do potential outcomes and causal diagrams differ?	11
	§ 9	How do you write a causal claim down?	12
	Figure 2	An identification record	13
	§ 10	Where does the evidence on causal inference stop?	14
		References	15

What is causal inference?

§ 1 of 10

Causal inference is the estimation of what a treatment changed. Gangl (2010)¹, following Heckman, describes it as “the attempt to recover causal parameters that describe social processes of interest from empirically observed data”. The treatment can be a programme, a price, a campaign or an event; the causal effect is the difference between the outcome a unit had and the outcome it would have had under the alternative.

¹ Gangl, M. (2010). Causal inference in sociological research. *Annual Review of Sociology*, 36, 21-47.

Rubin (1974)⁴ gives the everyday version. The causal effect of one treatment over another, for a particular unit and an interval of time, is the difference between what would have happened under each: “If an hour ago I had taken two aspirins instead of just a glass of water, my headache would now be gone”. The difficulty is inside the definition: “we can never observe both $y(E)$ and $y(C)$ ”, since no one can go back in time and give the same unit the other treatment. Rosenbaum and Rubin (1983)³ draw the consequence. “Since each unit receives only one treatment”, one of the two outcomes is always missing, and so “In a sense, estimating the causal effects of treatments is a missing data problem”.

³ Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.

⁴ Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701.

Gangl (2010)¹ reports that “the counterfactual, Rubin, or potential outcomes model of causality has, over the past three decades, become the standard conceptual tool” for defining causal effects and understanding the identification problem. The unit effects it defines “are therefore unobservable in principle but need to be estimated from empirically observable data under auxiliary identifying assumptions”.

In this guide, causal inference means estimating, for a named population, the difference between the outcomes it had and the outcomes it would have had under a named alternative, and stating the assumptions that let an observed comparison stand in for the outcomes nobody observed. The data supply the comparison; the design and its assumptions are what make it causal.

Several guides on this site take one part of that problem. The [incrementality guide](#) applies it to the revenue a commercial action added, and the A/B test guide covers the randomized experiment. The guides to [selection bias](#), interference and parallel trends each take one assumption that can fail, and the external validity guide asks whether an estimate travels to another target. This page sets out the frame they share.

Why is correlation not causation?

§ 2 of 10

Because the units that received a treatment usually differed from the others before it. Gangl (2010)¹ shows that “the straightforward comparison of average outcomes in the two groups (sometimes called the naive estimator) will be a biased estimator of the ATE parameter”, the average treatment effect. Two terms make up the bias: the groups “might have seen different baseline outcomes even absent the intervention”, and the effect may differ between those who took the treatment and those who did not. He names “selection bias, or heterogeneity, and endogeneity, or self-selection bias, as the two principal sources of bias when recovering any causal parameter from empirical data”. His selection bias is a difference in baseline outcomes between the groups, close to what Pearl, below, calls confounding bias; the selection bias guide on this site uses the term for distortion from the rule that decides which units enter the data.

Pearl (1995)² opens with an example from Cochran: a study of soil fumigants and oat yields in which “farmers insist on deciding for themselves which plots are to be fumigated”, and their choice depends on last year’s eelworm population, which is strongly correlated with this year’s. “Thus we have a classical case of confounding bias, which interferes with the assessment of treatment effects, regardless of sample size.” More data from the same comparison does not remove the bias.

² Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4), 669-688.

The same structure appears in commercial data, on this guide’s reading. Accounts that adopt a feature, join a training programme or receive a discount are often chosen, or choose themselves, for reasons tied to the outcome: they were already growing, already engaged or already at risk. A dashboard that compares their revenue with everyone else’s measures that selection together with any effect.

Gangl (2010)¹ does not reject regression on this account: “whereas correlation clearly is not causation, regression models do identify the

causal effect of interest if treatment status and outcomes are unconfounded”. What makes an estimate causal is the design and the assumptions behind it, not the method that computes it, although the method can change which average it reports.

Why do randomized experiments come first?

§ 3 of 10

Because assignment by chance cuts the link between who is treated and how they would have fared. “Randomized experiments in particular eliminate bias”, Gangl (2010)¹ writes, because they combine an active manipulation of the treatment with random assignment, which balances expected outcomes across the conditions. Rubin (1974)⁴ concludes “that randomization should be employed whenever possible but that the use of carefully controlled nonrandomized data to estimate causal effects is a reasonable and necessary procedure in many cases”.

Rubin’s⁴ comparison is careful. For variables not explicitly controlled, “a randomized study leaves the investigator in a more comfortable position than does a nonrandomized study”. Three points hold for both kinds of study. Any adjustment depends in part on whether the model behind it is right. “We can never know that all causally relevant prior variables that systematically differ in the E and C trials have been controlled”. And irrelevant prior variables have to be ignored even when they differ between the groups; otherwise, by searching long enough for the right ones, “we can obtain any estimate desired”.

An experiment is not a free pass. Even with full randomization, Gangl (2010)¹ notes, identification needs the stable unit treatment value assumption, which requires that outcomes not depend on assignment beyond a unit’s own treatment, “thus ruling out Hawthorne or John Henry effects” as well as “social interactions, information diffusion, norm formation” and other wider effects. He also warns that experimental estimates “may be difficult to generalize if either the study population or the administered treatments cannot be considered representative of some larger context”. Rubin (1974)⁴ puts the trade-off plainly: “Typically, nonrandomized studies have more representative trials than experiments since these are often conducted in constrained environments.”

How can observational data support a causal claim?

§ 4 of 10

Only under an assumption the data cannot prove. Gangl (2010)¹ states it as conditional independence: the observed covariates are a sufficient set of joint causes of the treatment and the outcome, “so that, conditional on covariates, variation in D is as good as randomly assigned”. Rosenbaum and Rubin³ (1983) give the stronger form, strongly ignorable treatment assignment: given the covariates, the outcomes are independent of the assignment, and every unit could have received either treatment. In a randomized trial with properly collected data, the first part is known to hold. “This condition is usually not known to hold in a nonrandomized experiment.”

The propensity score is their tool for using that assumption. “The propensity score is the conditional probability of assignment to a particular treatment given a vector of observed covariates”, and “adjustment for the scalar propensity score is sufficient to remove bias due to all observed covariates”. The word observed is the limit: the score balances what was measured and nothing else. Outside a randomized trial “the propensity score function is almost always unknown” and may be estimated, for example with a logit model.

Which covariates belong in the adjustment is a question of theory, not of fit. Gangl (2010)¹ writes that the conditional independence assumption “is not a primarily statistical, but rather a fundamentally theoretical statement about the joint causes X of D and Y ”. What is needed is not every joint cause of treatment and outcome but “a minimally sufficient set of covariates” that breaks the dependence between them. Pearl (1995)² supplies a graphical test for such a set, the back-door criterion, and states that it “is equivalent to the ignorability condition of Rosenbaum & Rubin (1983)³”.

Adding the wrong variable can do harm. Gangl (2010)¹ warns that controlling for a variable can “put the causal interpretation in jeopardy” if it “is an intervening mechanism that transmits the effect of D on Y ”, because “it is only the total effect of D on Y that has straightforward causal content”. He also warns, after Pearl, against “conditioning on collider variables”, which are typically meant to proxy an unobserved factor, because doing so “may inadvertently induce rather than alleviate dependency between treatment status and outcomes”. In a commercial reading, which is this guide’s, an analysis of whether an onboarding programme raised renewal should not control for later product usage if the programme works by raising usage.

What do natural experiments and repeated observations add?

§ 5 of 10

Natural experiments supply variation that the units did not choose; repeated observations let each unit serve as its own comparison. Natural experiments, in Gangl's (2010)¹ account, "all share the idea of exploiting some external event or institutional condition that creates exogenous variation in the social process of interest": a shock or a change in a programme, or, in a regression discontinuity design, a known "institutional rule to assign or withhold treatment that creates the relevant variation". The covariates must still account for other causes of change at the same time and for differences between the populations compared.

With repeated observations of the same units, fixed-effects and "difference-in-differences estimators may again be employed to condition on time-invariant unobservables", which weakens the conditional independence assumption. The price is two assumptions. The conditional independence that remains, Gangl notes, "is still a strong one": it rules out units that take the treatment because they correctly expect it to pay. And the fixed-effects estimator "rests on the assumption of common (parallel) time trends across groups", the subject of the [parallel trends guide](#), and with heterogeneous effects "FE estimates may have little [external validity](#)", because only the units whose treatment changed inside the window contribute.

An instrumental variable is a variable "that predicts treatment status D , but not outcomes" except through the treatment. Gangl (2010)¹ is direct about what it can deliver: "IV does not reliably identify the average treatment effect of D unless the" effect is the same for every unit. Under a monotonicity condition it identifies the local average treatment effect, for "those units i that complied with treatment D because of the instrument Z ". Because "LATE is defined by the specific instrument Z available to the researcher", two valid instruments can give two different answers, and "reliance on weak instruments undermines the statistical consistency and efficiency of the estimator".

Figure 1 Five routes to a causal estimate, and what each must assume

Each route builds the missing comparison a different way. Each must assume something the data cannot show, and they need not estimate the same effect.

Route	Where the comparison comes from	What it must assume	What it estimates	What can break it
Randomized experiment	A known chance mechanism assigns the treatment	Stable unit values: no unit's outcome depends on another's treatment or on being studied	The average treatment effect in the study population; with non-compliance, only the effect of assignment (intention-to-treat)	Spillover between units, Hawthorne or John Henry effects, an unrepresentative study population
Adjustment for observed covariates	Units alike on measured covariates, or on the propensity score, are compared	Conditional independence: the covariates form a sufficient set of joint causes of treatment and outcome (not necessarily all of them), and every unit could have received either treatment	The average effect, or the effect on the treated, where the groups overlap	An unmeasured joint cause the covariates do not account for; adjusting for a mediator or a collider

Route	Where the comparison comes from	What it must assume	What it estimates	What can break it
Natural experiment	An external event or institutional rule changes conditions for some units and not for others	Given the covariates, the event or rule is as good as randomly assigned: the covariates account for other causes of change at the same time and for differences between the groups it separates	The effect of the event or rule itself; using it to estimate the effect of a treatment it shifts is the instrumental-variable route	Change that happens at the same time; groups that already differed; few events that inform the question
Fixed effects and difference-in-differences	The same units or groups observed before and after	Unmeasured differences are constant over time, and the groups share a common trend	The effect for units whose treatment changed in the window	Diverging trends; units that time the treatment on its expected payoff; a selective set of units that changed

Route	Where the comparison comes from	What it must assume	What it estimates	What can break it
Instrumental variable	A variable that moves the treatment but not the outcome directly	The instrument is as good as randomly assigned and moves the treatment; it reaches the outcome only through the treatment; units respond to it in one direction	The local effect among units the instrument moved; another instrument can move other units and give another local effect	A weak instrument; a second path to the outcome; an instrument that shares a cause with the outcome

Read the third column first. The route is only as strong as the assumption it cannot fully test.

Source Author's synthesis of Rosenbaum and Rubin (1983), pp. 41, 43; Gangl (2010), pp. 24-29, 33-38; Rubin (1974), p. 699; and Pearl (1995), p. 671. The rows summarise the sources' conditions in the author's words; no estimate is reported.

When the covariates cannot justify conditional independence in full, Gangl (2010)¹ points to sensitivity analysis, which tries “to understand the extent of remaining bias in the empirical estimates under plausible assumptions about the degree of unmeasured confounding in the data”. Such analyses, he adds, “are most meaningful if the range of simulated values has empirical content”, for example when the likely size of confounding can be compared with the effect of a known covariate.

Which assumptions travel with a causal claim?

§ 6 of 10

Three of them. The first and the third apply whatever the route; the second binds wherever covariates do the adjusting. The first is that each unit has one outcome per treatment. Rosenbaum and Rubin (1983)³ note that the potential outcomes structure fails when one unit's response depends on the treatment another unit receives, "as could happen if they compete for resources"; the assumption that it does not "has been called the stable unit-treatment value assumption". Where social interactions matter for outcomes, "unit effects are inherently undefined", Gangl (2010)¹ writes, and he gives two responses: "An obvious response to potential violations of SUTVA is to move the analysis to a more aggregate level", such as an organization or a local labour market, or to accept that "standard treatment effect estimates are then to be considered as local", that is, tied to their time and situation. On this guide's reading, sales representatives who share a territory, or accounts that compare offers with each other, are commercial cases of the problem the interference guide treats.

The second is overlap, which Rosenbaum and Rubin (1983)³ build into strong ignorability. In a randomized experiment, "every unit in the population has a chance of receiving each treatment"; an adjusted comparison needs the same at every value of the covariates. Where some kinds of unit are always or never treated, the data hold no comparison for them: Gangl (2010)¹ notes that "matching cannot estimate a treatment effect in areas where covariate distributions do not overlap for the treatment and control group", and the estimate "only applies to the population defined over the common covariate support".

The third is that the assumptions are stated and argued, not left implicit. Rubin (1974)⁴ advises that "the investigator should think hard about variables besides the treatment that may causally affect Y and plan in advance how to control for the important ones", and that "it is important to indicate the extent to which the assumptions of subjective randomization and subjective random sampling can be believed", his terms for treating the assignment and the sample as if they were random.

Which effect is being estimated?

§ 7 of 10

There is more than one, and the choice comes first. Gangl (2010)¹ writes that with heterogeneous effects “there is no single number that would provide the causal effect of D on Y”, and that “Determining which causal parameter(s) to focus on in any given study thus becomes a first theoretical issue to be settled by the analyst”. The average treatment effect averages over everyone in the population. The average treatment effect on the treated averages over those who received the treatment, and “may be both easier to identify and theoretically more informative”. The local average treatment effect covers only those an instrument moved.

An average can mislead even when it is correct. Rubin (1974)⁴ notes that if one individual effect is very large and the rest are small, “the average causal effect may be substantially larger than all but one of the individual causal effects”. In a commercial reading, the effect of a discount on the accounts that were given it is a different number from its effect if it were offered to all accounts, and neither describes the typical account if a few large ones drive the result.

How do potential outcomes and causal diagrams differ?

§ 8 of 10

On Pearl’s² account, they state the same assumptions in two languages; he reads the equations behind a diagram as counterfactual relations, “forming a clear correspondence between causal diagrams and Rubin’s⁴ model of potential outcome”. The potential outcomes model writes the assumptions as conditions on unobserved outcomes, such as conditional independence. Pearl (1995) writes them as a diagram; in his framework “diagrams are queried to determine if the assumptions available are sufficient for identifying causal effects from nonexperimental data”. The diagram represents “the investigator’s understanding of the major causal influences among measurable quantities in the domain”, and his back-door criterion “can be applied directly to the causal diagram”. Gangl (2010)¹ draws on Pearl’s later analysis for the choice of covariates, including the warning about colliders.

Pearl is plain about the limit: “the results must rest on the causal assumptions shown in the graph, and that these cannot usually be tested in observational studies”. Some can at least be refuted: he notes that some, “most notably those associated with instrumental variables”, are subject to falsification tests, shown in related papers. Gangl (2010)¹ adds that an instrument’s exclusion restriction itself “is

untestable in principle and needs to be justified theoretically”. Pearl’s² case for the diagram is that “any causal inferences from observational studies must ultimately rely on some kind of causal assumptions”, and the diagram offers “an effective language for making those assumptions precise and explicit”.

How do you write a causal claim down?

§ 9 of 10

In two halves. The first says what the claim is about. The uplift specification essay gives a nine-field sheet for that half, written for marketing lift claims: unit, treatment, counterfactual, estimand, interference, time window, outcome, measurement and decision. Its release gates already ask for the assignment rule; the second half asks whether that rule supports the comparison, and says why the comparison is causal: how the units came to be treated, which route builds the comparison, what that route assumes, what was done to probe each assumption, and which statement survives. It follows Rubin’s⁴ advice to name the variables that may affect the outcome, plan their control in advance and say how far the assumptions can be believed, and Gangl’s¹ point that the effect of interest is chosen before the estimate.

Figure 2 An identification record

Beside the claim, write how units were assigned, which route builds the comparison, what it assumes, how each assumption was probed and which statement survives.

Claim	Assignment	Route	Assumptions	Probes	Supported statement
As in its specification: unit, estimand, outcome, window.	By chance, by choice or by a rule?	Which of the five routes in Figure 1 builds the comparison?	What must hold for the comparison to stand in for the missing outcome?	Which balance check, falsification test or sensitivity analysis was run?	Average, on the treated, local, ITT, or association only?

If units chose the treatment and the assumptions field is empty, the supported statement is an association.

One row per claim, kept beside its specification. The probes field is where an untested assumption shows.

Source Author's decision worksheet. The fields follow Rubin (1974), p. 700, on naming and controlling other causes and stating how far the assumptions can be believed; Gangl (2010), pp. 23-29, 32 and 36-37, on estimands, identification, sensitivity analysis and local effects; Rosenbaum and Rubin (1983), p. 43, on ignorability; and Pearl (1995), p. 684, on falsification tests. The fields are reader inputs; no company data is supplied.

A worked illustration, with no company data: a team sees that accounts which completed an onboarding programme renewed more often, and wants to credit the programme. The claim is the effect of completing the programme on renewal at twelve months, for the accounts that completed it. The assignment field says by choice, so the raw gap carries the first source of bias Gangl names, a difference in baseline outcomes: accounts that chose to complete would probably have renewed more often anyway. His second source, a difference in gains, matters for the average over all accounts; for the effect on those treated, he notes, “treatment effects on the treated will generally be identified under somewhat weaker conditions”, “although the practical utility of this result is probably limited”, and the line between the two sources “may also blur in practice”. Two designs would help. Adjusting for usage, size and tenure measured before onboarding rests on conditional independence and needs a sensitivity analysis. Randomizing the invitation instead gives the effect of the invitation,

and, used as an instrument, a local effect for the accounts the invitation moved to complete. Unless the programme is closed to accounts without an invitation, they are not all the completers, and “the population of compliers cannot actually be identified in the sample data”. Gangl notes that the local effect “identifies the complier average causal effect (CACE), whereas the net effect of the manipulation Z corresponds to the intention-to-treat (ITT) effect”. Either way, the claim field changes to the effect the design delivers. Until one of them is run, the supported statement is an association.

Where does the evidence on causal inference stop?

§ 10 of 10

At method. The four sources are methodological: Rubin (1974)⁴ writes for educational and psychological research, Rosenbaum and Rubin (1983) and Pearl (1995)² for statistics, with examples from medicine and agriculture, and Gangl (2010)¹ reviews the literature for sociology. None estimates a commercial effect. Gangl’s own conclusion favours design: “an emphasis on the primacy of research design seems the more natural starting point for many, if not most, attempts to recover causal parameters in sociology”. He adds that (quasi-)experimental methods “inevitably identify only selected causal parameters for specific populations and specific interventions”, so they complement observational analysis rather than replace it, and that, while “causal inference will inevitably remain a formidable challenge”, the prospects for causal analysis in sociology “may thus be better than its current reputation would suggest”. Pearl’s identification analysis also assumes that the sample is so large that sampling variability can be ignored, so an identified effect still needs its confidence interval.

This guide does not cover estimation in detail, statistical inference for an estimate, or methods published after 2010. The commercial examples are the author’s reasoning and have not been tested.

 End of essay

References

- 1 Gangl, M. (2010). Causal inference in sociological research. *Annual Review of Sociology*, 36, 21-47.
<https://doi.org/10.1146/annurev.soc.012809.102702>
- 2 Pearl, J. (1995). Causal diagrams for empirical research. *Biometrika*, 82(4), 669-688. <https://doi.org/10.1093/biomet/82.4.669>
- 3 Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
<https://doi.org/10.1093/biomet/70.1.41>
- 4 Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688-701. <https://doi.org/10.1037/h0037350>

How to cite this essay

Isoglu, S. (2026, October 18). What is causal inference? Sinan Isoglu.
<https://isoglu.com/insights/journal/what-is-causal-inference/>
Online version: <https://isoglu.com/insights/journal/what-is-causal-inference/>

Sinan Isoglu, MBA (Quantic)

Commercial growth leader, lecturer and doctoral researcher
© 2026 Sinan Isoglu · isoglu.com · All rights reserved