Vol. 11 (2026): Proceedings of the 2nd International Conference on Social Sciences and Cultural Studies (ICSSCS 2026)
ISSN: 3105-0522 P-ISSN：3105-0514
DOI: https://doi.org/10.70267/icsscs.20260111

Credibility Methods in Causal Inference

Yawen Zhan

Institute for Economic and Social Research, Jinan University, Guangzhou 510632, China

**Corresponding author: Yawen Zhan*

Abstract

How to credibly investigate causal relationships is a central issue in contemporary econometrics. This paper systematically reviews the logical chain of modern causal inference: from the paradigm shift of the credibility revolution, to the counterfactual framework for defining causality, to the assumptions and limitations of OLS regression, then to the concept of identification as the bridge connecting data to truth, and finally to the art of research design. At each stage, a study on the “Citation Trap” serves as a running case example to concretize abstract methodological principles. The analysis shows that a credible causal inference study begins with a clear “what if” question, approximates the ignorability assumption through control variables and fixed effects, and constructs a complete chain of evidence by ruling out alternative explanations and testing mechanism implications. The conclusion is that the essence of econometrics lies not in complex models or techniques, but in rigorous research design that makes assumptions plausible and conclusions robust to scrutiny.

Keywords

causality, credibility revolution, counterfactual framework, identification, research design

1. Introduction

How to distinguish correlation from causation? How to identify true causal effects from observational data? These are core methodological questions in economics and the broader social sciences. Everyday claims such as “smokers live longer,” “college graduates earn more,” or “major hospitals have higher mortality rates” essentially stem from conflating correlation with causation—smoking and longevity may merely coexist by chance, while higher earnings among college graduates or higher mortality in major hospitals may be confounded by individual ability or case-mix factors. Accurately identifying causal relationships is crucial for the validity of scientific research and the soundness of public policy.

Angrist and Pischke (2010) characterize the development of empirical research over the past nearly thirty years as the “Credibility Revolution”. Its core message is a shift in focus from complex model specification to careful research design, and from attempting to “explain the world” to answering specific causal questions. Within this framework, researchers no longer settle for reporting associations between variables; instead, through randomized experiments or quasi-experimental designs, they approximate counterfactual states to identify clean causal effects. Angrist and Pischke (2009) further systematize identification strategies such as instrumental variables, difference-in-differences, and regression discontinuity, providing methodological guidance for applied research.

Although the methodological system of causal inference has matured, existing literature largely focuses on introducing and applying specific methods, with relatively little systematic review of the logical structure underlying the credibility revolution. How to connect core concepts such as the counterfactual framework, identification strategies, and research design into a coherent logical chain, and how to integrate abstract methodological principles with concrete research practice, remain questions worthy of deeper exploration.

This paper attempts to fill this gap. The structure is as follows: first, starting from the credibility revolution, it elaborates the paradigm shift in modern empirical research; second, it introduces the counterfactual framework to clearly define causal effects; third, it analyzes the assumptions and limitations of ordinary least squares (OLS) regression, revealing its inherent constraints as a basic tool; fourth, it explains the concept of identification, clarifying the logic of inference from data to truth; finally, centering on research design, it demonstrates how to build credible evidence chains by narrowing the comparison scope, conducting robustness checks, testing mechanism hypotheses, and ruling out alternative explanations. At each stage, Ma et al.'s (2019) study on the "Citation Trap" serves as a running case example throughout, concretizing the abstract methodological principles. That study focuses on a clear causal question—"Does publication timing affect paper citations?"—faces typical identification challenges, and enhances the credibility of its conclusions through systematic research design.

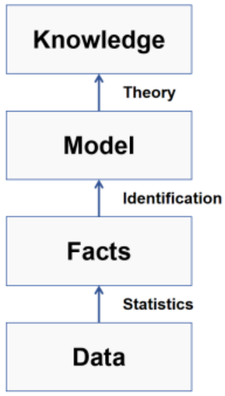
This paper aims to outline the methodological landscape of modern causal inference for beginners and to provide methodological reference for applied researchers. Its expected contributions are: first, systematically integrating core concepts such as the credibility revolution, counterfactual framework, identification strategies, and research design into a complete logical chain; second, combining abstract principles with concrete practice through a case study to improve operationalizability; third, emphasizing the central role of research design—rather than model complexity—in causal inference, offering methodological guidance for empirical research.

2. The Credibility Revolution: Paradigm Shift in Modern Empirical Research

This paradigm shift can be understood by comparing two classic studies. Summers and Wolfe (1977) attempted to "explain" student performance by including all variables in a regression model and interpreting coefficients indiscriminately. In contrast, Dale and Krueger (2002) focused on a specific causal question—does attending an elite university affect earnings?—and clearly distinguished treatment from control variables, eliminating selection bias through research design. This comparison reveals the core turn in modern empirical research: from explaining the world to answering specific causal questions, and from relying on complex models to relying on careful design.

Figure 1 illustrates the logical chain of empirical research. The framework shows that the essence of the credibility revolution lies in strengthening the inference chain from data to knowledge, making research design a solid bridge connecting theoretical models and empirical facts.

Figure 1: Empirical research logic map



Having understood this paradigm shift, the next question is: when we speak of "causality," what exactly are we talking about?

3. Counterfactual Framework and Causal Definition

Following the discussion of the paradigm shift, a fundamental question arises: what exactly is “causality”? Different answers to this question constitute two major schools in causal inference theory.

The theoretical foundations of causal inference can be traced to two philosophical traditions. The first views causality as an inherent property of the world, reasoning from observed outcomes to ask “why did this happen?” and seeking the causes behind effects (Causes of Effects). The second views causality as a comparison between alternative world states, reasoning from possible causes to ask “what if this factor were changed?” and identifying the effect of the cause by comparing treated and untreated states (Effects of Causes). Holland (1986), in systematically expounding Rubin’s (1974) causal model, analyzed the differences between these perspectives in depth and established the potential outcomes framework as the unified language of causal inference. The core differences between the two schools are summarized in Table 1.

Table 1: Comparison of the two major schools of thought on the definition of cause and effect

	Family 1	Family 2
School 1	Regularity	Counterfactuals
School 2	Mechanisms	Manipulation
What is causality?	Property of the world	Comparison between worlds
Direction	Top-down	Bottom-up
Question	"Why?"	"What if?"
Focus	Causes of Effects	Effects of causes
Model	Structural model	Potential outcomes

Modern econometrics mainly adopts the second perspective, with its theoretical framework established by Rubin’s (1974) Potential Outcomes Framework. The core idea can be summarized as “same unit, single treatment, contrasting outcomes.” For an individual *i*, the causal effect is defined as the difference between the potential outcome under treatment and under no treatment:

$$\tau_i = Y_i(1) - Y_i(0)$$

where τ_i is the causal effect for individual *i*; $Y_i(1)$ is the outcome if treated; $Y_i(0)$ is the outcome if untreated (Imbens & Rubin, 2015; Rubin, 1974).

This definition clearly reveals the fundamental problem of causal inference: for any individual, we can never observe both $Y_i(1)$ and $Y_i(0)$ simultaneously (Holland, 1986; Rubin, 2007). This dilemma forces researchers to seek alternatives, the most basic of which is linear regression. However, whether regression can identify causal effects depends entirely on whether its underlying assumptions hold.

Taking Ma et al.’s (2019) “Citation Trap” study as an example, the causal question can be clearly expressed as a counterfactual: if a paper published at the end of the year had instead been published at the beginning, how much would its citations change? For any paper *i*, $Y_i(1)$ represents citations if published at year-end, $Y_i(0)$ if published at year-start; the difference is the causal effect of publication timing. This case intuitively illustrates the application of the counterfactual framework in concrete research.

4. Assumptions and Limitations of Linear Regression Models

Faced with the fundamental problem of causal inference, ordinary least squares (OLS) regression is the most commonly used analytical tool. By minimizing the sum of squared residuals, it provides the best linear unbiased estimator (Wooldridge, 2010). In a simple regression, the OLS estimator of the slope coefficient is:

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \tag{1}$$

where $\widehat{\beta}_1$ is the estimated slope; x_i are observed values of the explanatory variable; $\bar{x}$ is its sample mean; y_i are observed values of the dependent variable; $\bar{y}$ is its sample mean.

However, OLS is merely a computational tool; whether its estimates have causal meaning depends entirely on the validity of its assumptions. To identify causal effects, OLS requires the conditional independence assumption $E[u|X]=0$, meaning that after controlling for all observable variables, the treatment variable *X* is uncorrelated with the error term *u*. This assumption is often difficult to fully satisfy in empirical research.

When the conditional independence assumption holds, the OLS estimator has two important properties: unbiasedness and consistency (Wooldridge, 2010). Unbiasedness means that, given the sample size, the expected value of the estimator equals the true value; consistency means that as the sample size approaches infinity, the estimator converges to the true value. Both properties depend on the conditional independence assumption; if violated, estimates become biased.

The most common source of bias is omitted variable bias (OVB) (Wooldridge, 2010). If the true model is

$$Y_i=\beta_0+\beta_1 X_i+\gamma W_i+u_i \tag{2}$$

but the researcher estimates

$$Y_i=\beta_0+\beta_1 X_i+v_i \tag{3}$$

where $v_i=\gamma W_i+u_i$, and the omitted variable W is correlated with X, then the estimate of β_1 will capture both the direct effect of X and the indirect effect transmitted through the correlation between W and X, making it impossible to isolate the pure causal effect of X.

In Ma et al.'s (2019) "Citation Trap" study, the researchers first use OLS and approximate the conditional independence assumption by controlling for variables such as page length, number of authors, number of references, and author team ability—factors that may affect both publication timing and citations—to make papers published in different months more comparable. However, controlling only observables cannot address unobservable confounders—for example, journal editors may schedule lower-quality manuscripts for year-end publication, or high-caliber scholars may deliberately avoid year-end submissions. If these factors are unobservable and uncontrolled, OLS estimates may remain biased.

This limitation raises a deeper question: how can we determine whether observed differences reflect true causal effects or biases from unobservables? This is precisely what the concept of identification addresses in the next section.

5. Identification Strategies: From Observation to Inference

Identification is the critical link connecting observed data to causal parameters. Manski points out that inference from observational data to causal parameters must rely on a series of identifying assumptions (Manski, 1995). This can be illustrated with a classic example.

Suppose a researcher wants to know the true cheating rate θ at a school but can only observe students' self-reported cheating rates θ_s . Inference from θ_s to θ requires different identifying assumptions. Table 2 summarizes identification results under various assumptions.

Table 2: *Identification results under different assumptions*

Strategy	Assumption	Identification Result
Strategy 1	No assumption	Cannot determine θ (no information)
Strategy 2	Strong assumption (no one lies)	$\theta=\theta_s$(Point identification)
Strategy 3	Weak assumption (non-cheaters do not lie)	$\theta\geq\theta_s$(Partial identification)
Strategy 4	External information (archive data)	Closer approximation θ (point identification)
Strategy 5	Experimental design ("coin flip" reduces lying probability)	$\theta=2\theta_s$-1 (Point identification)

This sequence of analyses reveals a core insight: the quality of an identification strategy depends on the plausibility of assumptions and the richness of information (Angrist & Pischke, 2009; Manski, 1995).

In the "Citation Trap" study, researchers face a similar identification problem. They observe that "papers published at year-end receive fewer citations on average" (Ma et al., 2019, Table 2], but this observed difference could arise from two entirely different explanations:

Observed difference = Causal effect + Selection bias

As in the classic "college and earnings" example: higher wages among college graduates may result from college increasing ability (causal effect) or from higher-ability individuals being more likely to attend college (selection bias). Similarly, lower citations for year-end papers may stem from publication timing itself or from lower-quality papers being more likely scheduled for year-end.

Distinguishing these possibilities is precisely what research design must accomplish.

6. Research Design and Construction of the Evidence Chain

Modern econometrics advocates a “design-based approach” (Angrist & Pischke, 2010), which differs from the “model-based approach” by relying not on complex model assumptions but on research design itself to enhance the credibility of conclusions. Key features of this paradigm include: orienting toward answerable questions, grounding in useful data, placing assumptions at the identification rather than theoretical level, and tailoring design strategies to specific contexts (Angrist & Pischke, 2009).

Ma et al.’s “Citation Trap” study (Ma et al., 2019) exemplifies this paradigm perfectly. Rather than stopping at OLS results, the researchers construct a complete evidence chain.

6.1 Narrowing the Comparison Scope

The researchers introduce “journal-year” fixed effects, limiting comparisons to papers published in different months within the same journal and year (Ma et al., 2019, Table 2]. This design controls for differences in journal quality and eliminates year effects, conducting comparisons in a much “cleaner” environment—this is a concrete application of the principle of “narrowing individual differences.”

6.2 Conducting Robustness Checks

The researchers test different definitions of “year-end,” different sample selection criteria, and different model specifications, with results remaining stable (Ma et al., 2019, Tables 3-6]. This demonstrates the consistency of the OLS estimator—the estimates converge to a stable range across specifications.

6.3 Testing Mechanism Hypotheses

The researchers propose the “search trap” hypothesis: lower citations for year-end papers result from academic search engines’ default settings creating a shorter retrieval window (Ma et al., 2019, Table 2]. This hypothesis yields two testable implications: if true, the effect should strengthen with the spread of internet search; and top journals, with fixed readership and less search dependence, should show a weaker effect. Empirical results support both: the effect is insignificant in the 1956-1970 subsample but highly significant in 2001-2010; top journals exhibit significantly smaller effects than ordinary journals (Ma et al., 2019, Table 7]. This “if…then…” style of testing elevates the mechanism hypothesis from conjecture to evidence-supported explanation.

6.4 Ruling Out Alternative Explanations

The researchers systematically test two potential confounders. First, do high-caliber scholars deliberately avoid year-end submissions? Comparison of author team ability across months shows no significant differences; survey evidence also indicates no widespread habit of avoiding year-end (Ma et al., 2019), Table 8, Table A3]. Second, do editors preferentially schedule high-quality manuscripts for early-year publication? Statistics on Nobel laureates’ publication months show no concentration in early-year issues (Ma et al., 2019, Fig. 4]. These tests rule out the most common alternative explanations, enhancing the credibility of the causal inference.

Through this progressive layering of design elements, the researchers bring the observed “fact” progressively closer to the true “causal effect.” This is the complete path from the concept of identification to its realization through research design—first understanding the essence of identification, then using design to achieve it.

7. Results and Analysis

This paper systematically reviews the theoretical framework and core strategies of modern causal inference from a methodological perspective. The preceding discussion yields three levels of findings:

First, the way a research question is framed determines the inference path. The key contribution of the credibility revolution is shifting research orientation from vague "explaining the world" to clear "answering specific causal questions." The contrast between Summers and Wolfe (1977) and Dale and Krueger (2002) shows that, although both use regression, the former attempts to explain all variables while the latter focuses on one core causal question; the starting point of research design determines the credibility of conclusions.

Second, the strength of identifying assumptions affects the reliability of conclusions. The "student cheating rate" example shows that with no assumptions, no conclusion is possible (Strategy 1); strong assumptions (no one lies) enable point identification but lack realism (Strategy 2); weak assumptions (non-cheaters do not lie) yield only partial identification (Strategy 3); external information brings closer approximation (Strategy 4); experimental design achieves point identification under reasonable assumptions (Strategy 5). Thus, the quality of identification strategies depends on the plausibility of assumptions and the richness of information (Manski, 1995).

Third, the rigor of research design is the fundamental guarantee of credibility. Ma et al.'s (2019) "Citation Trap" study demonstrates how to build an evidence chain through progressive design: first, introducing journal-year fixed effects to narrow the comparison scope and eliminate journal quality and year effects; second, testing robustness across multiple specifications to show conclusions do not depend on particular definitions or samples; third, using time trends and journal heterogeneity to test the mechanism hypothesis, elevating the "search trap" from conjecture to evidence-supported explanation; finally, ruling out author strategy and editor preference as alternative explanations, strengthening causal credibility.

These three levels collectively indicate that the credibility of causal inference depends not on model complexity but on the rigor of research design.

8. Conclusion

Reviewing the logical structure of the paper, the methodological landscape of modern causal inference becomes clear. Its essence can be summarized in three points:

First, the essence of causal inference is comparing counterfactuals. Most causal questions should be reframed as clear "what if" questions: what would be different if the treatment were changed? This framing forces researchers to confront the fundamental problem of causal inference and seek solutions through research design.

Second, identification is the bridge from data to truth. Inference from observed "facts" to desired "truth" requires a series of assumptions to build the bridge. The quality of identification strategies depends on the plausibility of assumptions and the richness of information—"no assumptions, no conclusions."

Third, research design is the safeguard of credibility. A credible causal inference study must approximate the ignorability assumption through control variables, fixed effects, etc., and construct a complete evidence chain by ruling out alternative explanations and testing mechanism implications. The power of econometrics lies not in the complexity of models but in the rigor of research design.

A limitation of this paper, as a methodological review, is that it does not delve deeply into the applicability conditions and limitations of specific identification methods such as instrumental variables, difference-in-differences, and regression discontinuity. Future work could further explore selection criteria for these methods in different contexts and how to apply research design principles to a broader range of empirical problems.

For beginners, the greatest value of this methodological framework lies not in mastering any specific technique but in cultivating a critical mindset—when encountering any "research finding," one should ask: How was the question defined? What is the identification strategy? Are the key assumptions plausible? Have alternative explanations been ruled out? Has the mechanism hypothesis been tested? This way of thinking may be the most valuable gift left by the credibility revolution to learners.

References

Angrist, J. D., Pischke, J. S. The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics, Journal of Economic Perspectives. 2010, 24(2), 3-30. https://doi.org/10.1257/jep.24.2.3

Angrist, J. D., Pischke, J.-S. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton: Princeton University Press; 2009.

Ma, C., Li, Y., Guo, F., Si, K. The Citation Trap: Papers Published at Year-End Receive Systematically Fewer Citations, Journal of Economic Behavior and Organization. 2019, 166, 667-687. https://doi.org/10.1016/j.jebo.2019.08.007

Summers, A. A., Wolfe, B. L. Do Schools Make a Difference?, The American Economic Review. 1977, 67(4), 639-652.

Dale, S. B., Krueger, A. B. Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables, The Quarterly Journal of Economics. 2002, 117(4), 1491-1527. https://doi.org/10.1162/003355302320935089

Holland, P. W. Statistics and Causal Inference, Journal of the American Statistical Association. 1986, 81(396), 945-960. https://doi.org/10.1080/01621459.1986.10478354

Rubin, D. B. Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, Journal of Educational Psychology. 1974, 66(5), 688-701. https://doi.org/10.1037/h0037350

Imbens, G. W., Rubin, D. B. Causal Inference: For Statistics, Social, and Biomedical Sciences an Introduction. Cambridge: Cambridge University Press; 2015.

Rubin, D. B. Bayesian Inference for Causal Effects: The Role of Randomization, The Annals of Statistics. 1978, 6(1), 34-58. https://doi.org/10.1214/aos/1176344064

Wooldridge, J. M. Econometric Analysis of Cross Section and Panel Data. Cambridge: MIT Press; 2010.

Manski, C. F. Identification Problems in the Social Sciences. Cambridge: Harvard University Press; 1995.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access