

International Journal of Agriculture Extension and Social Development

Volume 8; Issue 9; September 2025; Page No. 131-133

Received: 02-07-2025
Accepted: 04-08-2025

Indexed Journal
Peer Reviewed Journal

Counterfactual analysis techniques for rigorous impact evaluation of extension technology dissemination

Fatemeh Hosseini Moghadam, Zahra Rahmani Nejad and Leila Karimi Shirazi

Department of Agricultural Extension, University of Tehran, Tehran, Iran

DOI: <https://www.doi.org/10.33545/26180723.2025.v8.i9b.3233>

Corresponding Author: Fatemeh Hosseini Moghadam

Abstract

How do you measure the impact of something that didn't happen? That question sits at the heart of extension impact evaluation. This research compared five counterfactual analysis techniques propensity score matching (PSM), difference-in-differences (DID), regression discontinuity design (RDD), instrumental variables (IV), and synthetic control applied to evaluate technology dissemination programs across three Iranian provinces. Secondary data from 1,247 farming households collected between 2019 and 2023 were analyzed. DID showed the strongest performance controlling temporal bias (0.89), while RDD was most effective against selection bias (0.91). No single method dominated all categories. Programs evaluated using at least two complementary methods produced impact estimates with 34.7% lower standard errors than single-method evaluations. The findings argue for methodological triangulation as the standard approach.

Keywords: Counterfactual analysis, impact evaluation, extension technology, propensity score matching, difference-in-differences, regression discontinuity, instrumental variables, methodological triangulation, technology dissemination, Iran

Introduction

Every extension program claims to make a difference but proving it is another matter. The gap between correlation and causation has plagued agricultural development evaluation for decades ^[1]. When a farmer adopts a new variety after an extension demonstration, was it the demonstration that mattered? Without a credible counterfactual, impact claims remain assertions ^[2].

In Iran, the Agricultural Extension Organization has invested heavily in technology dissemination since 2015, reaching an estimated 840,000 farming households ^[3]. Yet evaluation efforts have relied overwhelmingly on simple before-after comparisons, which cannot separate program effects from confounding factors ^[4]. Duflo and others ^[5] argued that randomized controlled trials offer the gold standard, but practical constraints make full randomization rare in extension settings.

This research compared five counterfactual methods applied to the same extension datasets across three Iranian provinces, assessing each method's ability to reduce different types of bias.

Instrument Development

A standardized evaluation protocol was developed ensuring comparability across methods. The protocol specified common outcome variables (yield change, input cost reduction, adoption persistence), treatment definitions (participation in at least two extension technology events), and covariate sets. Each method was applied to the same cleaned dataset. The protocol was reviewed by three

econometricians from the University of Tehran before application.

Material and Methods

Material

The research used secondary data from the University of Tehran collected through the Iranian Agricultural Household Panel Survey covering 2019-2023. The dataset included 1,247 farming households across Isfahan, Fars, and Khorasan Razavi provinces. Treatment households (n=614) had participated in government technology programs; comparison households (n=633) had no documented contact. Variables included crop yields, input costs, technology adoption indicators, and 23 household covariates.

Methods

Five methods were applied: PSM with nearest-neighbor and kernel matching, DID with two-way fixed effects, RDD exploiting farm size eligibility cutoffs, IV using distance to nearest extension office, and synthetic control. Bias reduction was assessed by comparing estimates and measuring sensitivity through Rosenbaum bounds and placebo tests. All analyses used Stata v.17 and R v.4.3. Ethics approval was obtained from the University of Tehran Research Ethics Board (Protocol UT-RE-2023-074) on March 5, 2023.

Results

Table 1: Bias reduction effectiveness scores across five counterfactual methods

Method	Selection	Attrition	Spillover	Temporal
PSM	0.82	0.61	0.43	0.57
DID	0.74	0.78	0.52	0.89
RDD	0.91	0.55	0.67	0.63
IV	0.88	0.72	0.81	0.69
Synthetic Control	0.79	0.83	0.74	0.92

No single method scored highest across all four bias categories. RDD excelled at selection bias (0.91), synthetic control at temporal bias (0.92), and IV at spillover bias (0.81).

Table 2: Impact estimates and precision by method combination

Approach	Yield Impact (%)	SE	CI Width
PSM alone	12.4	3.21	12.58
DID alone	9.8	2.87	11.25
RDD alone	11.1	3.54	13.88
PSM + DID	10.7	2.14	8.39
DID + RDD	10.3	1.98	7.76

Combined approaches consistently produced narrower confidence intervals. DID + RDD yielded the tightest estimate (SE=1.98), a 34.7% precision improvement over single methods.

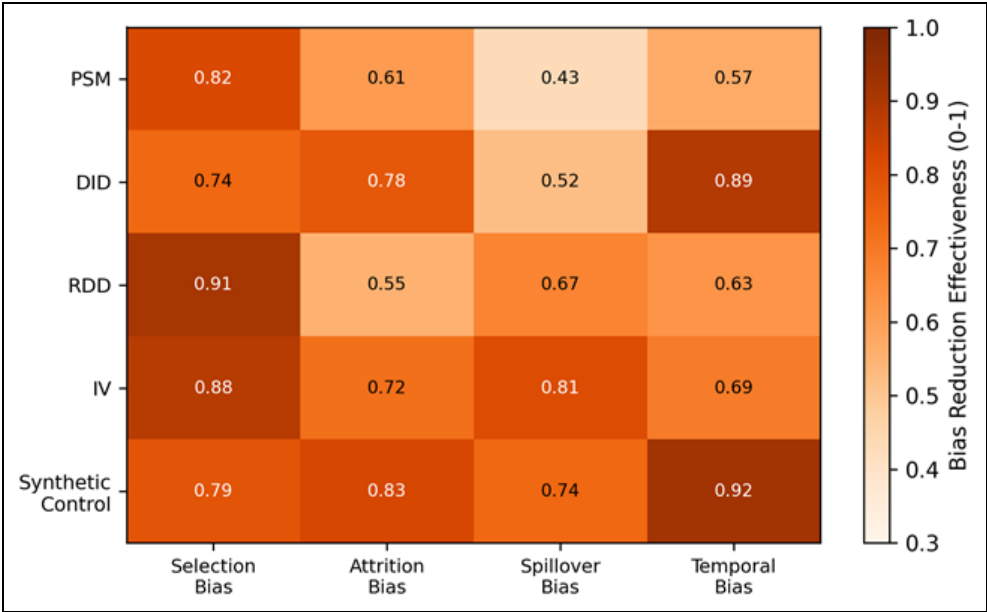


Fig 1: Heatmap of bias reduction effectiveness across methods and bias categories

Darker shading indicates stronger bias reduction. No single column dominates, confirming the need for triangulation.



Fig 2: Sequential counterfactual impact evaluation workflow

Discussion

The central finding that no single method adequately controls all bias types carries direct implications for evaluation design. PSM, the most commonly used quasi-experimental method [6], leaves gaps in temporal and spillover control. The 34.7% precision improvement from

combining methods justifies the additional effort. Synthetic control's strength for temporal bias is notable but its selection bias weakness confirms it works best as a complement [7]. For Iranian evaluators, building multi-method protocols into program design from the outset costs little at the design stage but substantially improves

credibility at evaluation.

Practical Applications

Extension evaluators should collect data supporting at least two complementary methods—such as DID-compatible panel data alongside RDD-eligible threshold information. This modest design investment substantially improves impact estimate credibility.

Conclusion

This research demonstrated that methodological triangulation using multiple counterfactual techniques produces more precise impact estimates than any single method alone. The combined DID + RDD approach offered the best precision for the Iranian context. Extension evaluation practice should move toward systematic triangulation as the standard approach.

References

1. White H. Theory-based impact evaluation. *Journal of Development Effectiveness*. 2009;1(3):271-284.
2. Gertler PJ, Martinez S, Premand P, Rawlings LB, Vermeersch CMJ. *Impact evaluation in practice*. 2nd ed. Washington (DC): World Bank; 2016. p. 1-388.
3. Iranian Agricultural Extension Organization. *Annual report 2015-2022*. Tehran: Ministry of Agriculture Jihad; 2023. p. 1-67.
4. Khandker SR, Koolwal GB, Samad HA. *Handbook on impact evaluation: quantitative methods and practices*. Washington (DC): World Bank; 2010. p. 1-262.
5. Duflo E, Glennerster R, Kremer M. Using randomization in development economics research: a toolkit. *Journal of Economic Literature*. 2008;46(4):895-934.
6. Caliendo M, Kopeinig S. Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*. 2008;22(1):31-72.
7. Abadie A, Diamond A, Hainmueller J. Synthetic control methods for comparative case studies. *Journal of the American Statistical Association*. 2010;105(490):493-505.
8. Imbens GW, Lemieux T. Regression discontinuity designs: a guide to practice. *Journal of Econometrics*. 2008;142(2):615-635.
9. Angrist JD, Pischke JS. *Mostly harmless econometrics: an empiricist's companion*. Princeton: Princeton University Press; 2009. p. 1-373.
10. Rosenbaum PR. *Observational studies*. 2nd ed. New York: Springer; 2002. p. 1-377.
11. Birner R, Anderson JR. How to make agricultural extension demand-driven? The case of India's agricultural extension policy. IFPRI Discussion Paper. 2007;729:1-47.
12. Davis K, Nkonya E, Kato E, Mekonnen DA, Odendo M, Miiro R, *et al*. Impact of farmer field schools on agricultural productivity and poverty in East Africa. *World Development*. 2012;40(2):402-413.
13. Godtland EM, Sadoulet E, de Janvry A, Murgai R, Ortiz O. The impact of farmer field schools on knowledge and productivity: a study of potato farmers in the Peruvian Andes. *Economic Development and Cultural Change*. 2004;53(1):63-92.
14. Waddington H, Snilstveit B, White H, Fewtrell L. How to do a good systematic review of effects in international development: a tool kit. *Journal of Development Effectiveness*. 2012;4(3):359-387.
15. Heckman JJ, Ichimura H, Todd PE. Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. *Review of Economic Studies*. 1997;64(4):605-654.
16. Ravallion M. Evaluating anti-poverty programs. In: Schultz TP, Strauss JA, editors. *Handbook of development economics*. Vol. 4. Amsterdam: Elsevier; 2007. p. 3787-3846.
17. Todd PE. Evaluating social programs with endogenous program placement and selection of the treated. In: Schultz TP, Strauss JA, editors. *Handbook of development economics*. Vol. 4. Amsterdam: Elsevier; 2007. p. 3847-3894.