ANNEX F

Supplement to PHREVO Framework Paper, Version 1.0
SSRN Abstract ID: 6614438 — DOI: 10.5281/zenodo.19666941

PHREVO Empirical Validation Protocol: Quasi-Experimental Longitudinal Study Design for Evaluating the Impact of PHREVO Implementation in Real Territories

RCT Protocol v1.0 — Research Protocol — Economic and Social Impact

Author

Andres Jimenez

Founder, PHREVO — Independent Research Initiative

Date
April 2026
Status
Proposal for funding and peer review — requires IRB approval before implementation
Audience
Philanthropic foundations, cooperation agencies, ethics committees, evaluation teams, academic collaborators
Pre-registration
Protocol to be registered at AEA RCT Registry and OSF (Open Science Framework) before data collection begins

Methodological Note: Why Quasi-Experimental Rather Than Pure RCT

In social sciences, the randomized controlled trial (RCT) is the gold standard for establishing causal inference. However, in territorial implementations of complex systems such as PHREVO, full randomization is often impracticable or ethically problematic for three reasons specific to this context.

First, territorial self-selection is inherent to PHREVO's design. PHREVO is only implemented where communities and local governments have explicitly requested it — consistent with the core principle that no territory can be assigned to PHREVO without its own institutional consent. Randomizing assignment would violate this principle.

Second, geographic separation between treatment and control units is often too large for standard RCT designs. Control territories may be 500 kilometers from

treatment territories, making them plausibly different on many dimensions that cannot be measured or controlled.

Third, the intervention itself (PHREVO implementation) is not a simple, discrete treatment — it is a complex, adaptive, multi-component system that interacts with local governance structures in ways that cannot be fully specified in advance.
This design therefore adopts a quasi-experimental approach combining three complementary identification strategies:
Propensity Score Matching (PSM) — paired assignment of similar territories to treatment and control conditions based on observed pre-treatment characteristics.
Difference-in-Differences (DiD) — the primary causal identification strategy, which uses within-territory variation over time and cross-territory variation to estimate the causal effect of PHREVO.
Interrupted Time Series (ITS) — for autonomous territories and cases where matching is not feasible, using each territory as its own control through pre-post comparison.
Cluster Randomization — where contextually possible (e.g., multiple neighborhoods within a single city), introducing genuine randomization at the cluster level.

The goal is not methodological purity but maximum credible evidence. A well-executed quasi-experimental study with transparent pre-registration, rigorous matching, and honest discussion of limitations produces more actionable knowledge than a poorly-implemented RCT.

Contents

F.1 Objectives and Hypotheses

F.1.1 General Objective

To evaluate the causal impact of PHREVO framework implementation (in minimum and complete intervention versions) on the multidimensional wellbeing of territories, measured through the PHREVO-Score and independently validated indicators, across a five-year longitudinal study.

F.1.2 Research Hypotheses

Null hypothesis (H0): PHREVO implementation produces no statistically significant effects on the multidimensional wellbeing of the territory compared with the absence of intervention (or with standard alternative interventions).

F.2 Study Design

F.2.1 Territory Selection Criteria

Inclusion Criteria

Exclusion Criteria

Territories where PHREVO is already being informally implemented (treatment contamination).
Territories with fewer than 5,000 inhabitants (insufficient statistical power).
Territories where the principal investigator (Andres Jimenez) has personal or financial conflicts of interest.

Target Sample

Minimum for statistical power: 6 territories (3 treatment + 3 control).
Ideal: 10 territories (5 treatment + 5 control).
With cluster randomization: 20 neighborhoods or districts within 2-3 cities (10 treatment + 10 control).

F.2.2 Assignment to Groups

Primary Design: Propensity Score Matching (PSM)

Step 1: Identify a candidate pool of territories meeting inclusion criteria (target N: 10-20 territories).

Step 2: Collect pre-treatment baseline variables:

PHREVO-Score (or its dimensions, where available from prior data).
GDP per capita or equivalent economic activity indicator.
Gini coefficient.
Multidimensional Poverty Index (MPI).
Population size and urban/rural composition.
Geographic characteristics (altitude, market access, biome).
Historical voting patterns (proxy for political orientation).

Step 3: Estimate propensity scores — the probability of each territory being selected for treatment, conditional on its baseline characteristics, using logistic regression.
Step 4: Match each treatment territory with the most similar control territory (nearest-neighbor matching 1:1 or 1:2).
Step 5: Verify balance between matched groups — standardized mean difference < 0.25 for all baseline variables after matching.

Secondary Design: Cluster Randomization (where feasible)

When a single city contains multiple comparable neighborhoods, cluster randomization is used:

Random assignment: 10 neighborhoods to treatment, 10 to control.
Unit of randomization: neighborhood or district.

Stratification by socioeconomic level to ensure balance on key variables.

Autonomous Territory Design (Type 2 Adopters)

Autonomous territories cannot be randomized or matched in the conventional sense. For these:

Interrupted Time Series (ITS) with the territory as its own control (baseline pre-intervention trend vs. post-intervention trend).

Synthetic control method — constructing a synthetic control territory as the weighted combination of other territories that best approximates the treated territory in the pre-intervention period.

F.2.3 Study Period and Phases

Total study duration: 5 years (60 months) for MIP; 7 years (84 months) for CIP.

Recommendation: begin with MIP — lower cost, lower legal friction, faster initial results. If 3-year results are positive, seek additional funding to extend to CIP in 1-2 territories.

F.3 The Intervention: What Does "Implementing PHREVO" Mean?

F.3.1 Minimum Intervention Package (MIP)

For territories where complete implementation is infeasible due to political, legal, or capacity constraints, the MIP preserves the central elements of PHREVO while minimizing logistical and political complexity.

Not included in MIP: PHREVO-Exchange (impact token market); multi-layer collective property (full PTIERRA); liquid governance with justice weights (full PGOB); Basic Impact Income (RIB). These are deferred to the CIP to manage cost and complexity.

F.3.2 Complete Intervention Package (CIP)

The CIP includes all MIP components plus:

CIP cost is significantly higher (RIB, Exchange development, legal support). Recommended for 1-2 territories with high institutional capacity and strong political will.

F.3.3 Implementation Fidelity

Territories falling below 60% fidelity on two consecutive measurements are reclassified as "intention to treat" (ITT) rather than "treatment received" (TOT) and analyzed separately. This distinction is pre-specified in the analysis plan to prevent selective reporting.

F.4 Indicators and Variables

F.4.1 Primary Dependent Variable

The territorial PHREVO-Score: the population-weighted average of individual PHREVO-Scores within the territory, combined with aggregate territorial indicators (ecological, governance, infrastructure).

Why this is the primary variable: it is the central metric of the system. If PHREVO implementation does not improve its own metric, the system fails by definition — and this is a result worth knowing. Score range: 0-1. Score calculation: performed by an independent evaluation team with no affiliation to PHREVO.

F.4.2 Secondary Dependent Variables

F.4.3 Control Variables

F.5 Data Collection

F.5.1 Data Sources

F.5.2 Collection Frequency

F.5.3 Sample Size and Statistical Power

Assumptions for power calculation:

Minimum detectable effect (MDE): improvement of 0.10 points on the 0-1 PHREVO-Score scale (e.g., from 0.50 to 0.60).

Within-territory standard deviation: 0.15 (preliminary estimate from comparable social wellbeing indices).

Number of territories: 6 (3 treatment + 3 control) as minimum; 10 as ideal.

Study period: 5 years, 6 measurement rounds.

Intraclass correlation (ICC): 0.05 (standard assumption for territorial clustering).

Significance level: alpha = 0.05.

Statistical power: 1-beta = 0.80.

Power calculation for DiD with clusters:

MDE = t(alpha/2, df) * sqrt[2 * sigma^2 * (1-rho) / (n*T)] * sqrt[1 + (m-1)*rho]

Where: t(alpha/2, df) = 2.0 (for small degrees of freedom); sigma^2 = 0.0225; rho = 0.05 (ICC); n = 6 (territories); T = 6 (time periods); m = 500 (survey households per territory per period).

Result: MDE = 0.09 — slightly below the 0.10 target. With 6 territories, the study can detect a 9-10% improvement in the Score at 80% power.

Recommendation: Increasing to 10 territories (5 and 5) reduces MDE to approximately 0.07 and achieves power above 90% — the preferred design if funding permits.

F.6 Analysis Strategy

F.6.1 Difference-in-Differences (DiD) — Primary Strategy

The DiD regression model:

Y(j,t) = alpha + beta * PHREVO(j,t) + gamma(j) + delta(t) + epsilon(j,t)

Where: Y(j,t) is the PHREVO-Score of territory j at time t; PHREVO(j,t) is a dummy variable equal to 1 if territory j received the intervention AND t is post-treatment; gamma(j) are territory fixed effects (capture time-invariant between-territory differences); delta(t) are time fixed effects (capture common time trends across all territories); epsilon(j,t) is the error term, clustered by territory for valid standard errors.

Beta is the DiD estimator — the average causal effect of PHREVO on the Score, net of pre-existing territory differences and common time trends.

Assumption verification:

Parallel trends pre-treatment: plot Y trajectories in the pre-intervention period; test the interaction (time x treatment group). If trends are not parallel, use DiD with territory-specific linear trends or event-study specification.

No contamination: verify that control territories have not informally adopted PHREVO practices through qualitative monitoring.

Stable composition: no mass migration that changes the population composition within territories.

No anticipatory effects: test for significant pre-trends in the treatment group before implementation begins.

F.6.2 Interrupted Time Series (ITS) — For Cases With Few Controls

ITS model for autonomous territories:

Y(t) = beta_0 + beta_1*T(t) + beta_2*P(t) + beta_3*(T(t)-T_0)*P(t) + epsilon(t)

Where: T(t) is time (months from series start); P(t) = 0 for pre-intervention, 1 for post-intervention; T_0 is the intervention point; beta_1 is the pre-intervention trend (change per month before PHREVO); beta_2 is the immediate level change at intervention; beta_3 is the slope change after intervention (sustained effect). All three are interpretable causal parameters when the parallel trends assumption is satisfied within the series.

F.6.3 Qualitative Process Analysis

To understand not only whether PHREVO works but how and why it works — or fails — three qualitative methods are applied:

F.7 Ethical Considerations

F.7.1 Governing Principles

F.7.2 Adverse Effects Protocol

F.7.3 Required Ethical Approvals

Institutional Review Board (IRB) approval from a recognized university (e.g., Universidad Nacional de Colombia, Universidad de Buenos Aires, LSE, MIT, or equivalent). The independent IRB requirement is non-negotiable — self-review by PHREVO is insufficient.

Local government permission (for municipalities) or community assembly decision (for autonomous territories), documented in formal minutes.

Individual informed consent for all survey and interview participants. Consent process adapted to local literacy levels, oral consent documented where written consent is not feasible.

F.8 Timeline and Budget

F.8.1 Five-Year Timeline

F.8.2 Budget Estimates (5 Years, 10 Territories)

Recommended starting point: MIP in 6 territories (3 treatment + 3 control), estimated cost $1.5M - $2M for 3 years. If 3-year results are positive, seek additional funding to extend to CIP and 10 territories.

F.8.3 Priority Funding Sources

F.9 Publication and Dissemination Plan

F.9.1 Pre-Registration and Transparency

Protocol pre-registration on: AEA RCT Registry and OSF (Open Science Framework) before any data collection begins. Pre-registration is binding for primary outcomes — secondary and exploratory analyses must be declared as such if not pre-specified.

Pre-analysis plan: all primary analyses specified before viewing outcome data, including exact regression specifications, handling of missing data, and criteria for ITT vs. TOT classification.

Open data (anonymized): public repository release at study end, or earlier if the primary results paper is published.

F.9.2 Publication Plan

F.9.3 Academic Co-Authorship Strategy

To maximize credibility, 1-2 independent academics (with no PHREVO affiliation) will be invited as co-authors on primary publications. Required profile: economist with impact evaluation experience (RCTs or DiD) in the Global South; preferably Latin American; no conflicts of interest with PHREVO (no prior consulting, no investments).

Target institutions: CEDLAS (Argentina), FAIR (Colombia), World Bank impact evaluation group, Inter-American Development Bank researchers, public university economics departments (UBA, UNAL, UNAM, USP).

F.10 Limitations

No impact study of this magnitude will be perfect. The strength of this design is its combination of quantitative rigor (DiD), qualitative depth (process analysis), and transparency (pre-registration, open data). Even with limitations, this will be the best available evidence about a post-capitalist system implemented in real territories.

F.11 Conclusion and Call for Funding

F.11.1 What This Study Is

This protocol presents a rigorous quasi-experimental design for the longitudinal empirical validation of PHREVO. It is not a pilot. It is a five-year study designed to produce peer-reviewed evidence on whether a post-capitalist economic framework can measurably improve multidimensional wellbeing in real territories — and under what conditions.

Without longitudinal data, PHREVO will remain vulnerable to the most important criticism any transformative framework faces: "theoretically interesting — but does it work?" This study exists to answer that question with the highest methodological standard currently feasible for complex territorial interventions.

F.11.2 What This Study Needs

This study requires four things simultaneously:

Philanthropic funding — from foundations committed to rigorous evidence about economic alternatives, not just advocacy or demonstration projects. The expected return is not financial; it is evidence for systemic transformation.

Academic collaboration — economists with impact evaluation expertise who bring methodological rigor and publication credibility that PHREVO alone cannot provide.

Cooperation agencies willing to consider PHREVO as a potential tool for their local development programs — and to contribute to the evidence base through co-financing.

Local governments willing to be pilot territories — with the understanding that the study covers PHREVO implementation costs and that results, positive or negative, will be published openly.

F.11.3 Immediate Next Steps

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic Control Methods for Comparative Case Studies. Journal of the American Statistical Association, 105(490), 493-505.

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

Banerjee, A., & Duflo, E. (2011). Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. PublicAffairs.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? Quarterly Journal of Economics, 119(1), 249-275.

Callaway, B., & Sant'Anna, P.H.C. (2021). Difference-in-Differences with Multiple Time Periods. Journal of Econometrics, 225(2), 200-230.

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.

Gertler, P., Martinez, S., Premand, P., Rawlings, L., & Vermeersch, C. (2016). Impact Evaluation in Practice, 2nd Edition. World Bank Group.

Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. Journal of Econometrics, 225(2), 254-277.

Heckman, J.J., & Vytlacil, E. (2007). Econometric Evaluation of Social Programs. In J.J. Heckman & E.E. Leamer (Eds.), Handbook of Econometrics, Vol. 6B. Elsevier.

Imbens, G.W., & Lemieux, T. (2008). Regression Discontinuity Designs: A Guide to Practice. Journal of Econometrics, 142(2), 615-635.

Jimenez, A. (2026). PHREVO: A Post-Capitalist Economic Architecture for the Global South. SSRN Working Paper 6614438. DOI: 10.5281/zenodo.19666941.

Lee, D.S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. Review of Economic Studies, 76(3), 1071-1102.

Ostrom, E. (1990). Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press.

Rabe-Hesketh, S., & Skrondal, A. (2012). Multilevel and Longitudinal Modeling Using Stata, 3rd Edition. Stata Press.

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.

White, H., & Sabarwal, S. (2014). Quasi-Experimental Design and Methods. UNICEF Office of Research — Methodological Briefs: Impact Evaluation No. 8.

PHREVO Framework Paper — Annex F — RCT Protocol v1.0 — April 2026

Supplement to SSRN 6614438 — DOI: 10.5281/zenodo.19666941This study is registered in the AEA RCT Registry. Unique identifying number: AEARCTR-0018502.

Andres Jimenez — hello@phrevo.earth — phrevo.org

Requires IRB approval before implementation. Pre-registration before data collection is mandatory.