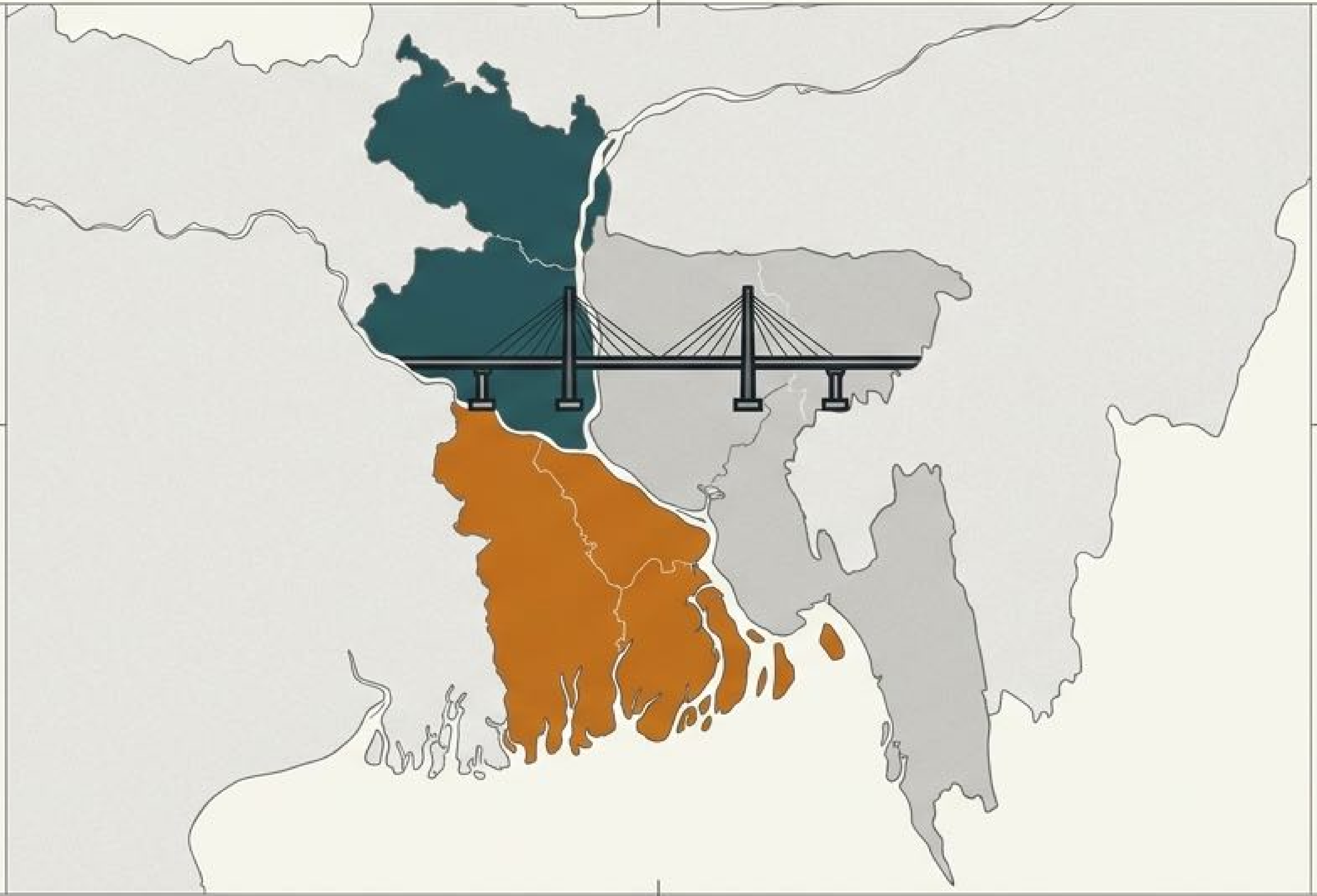
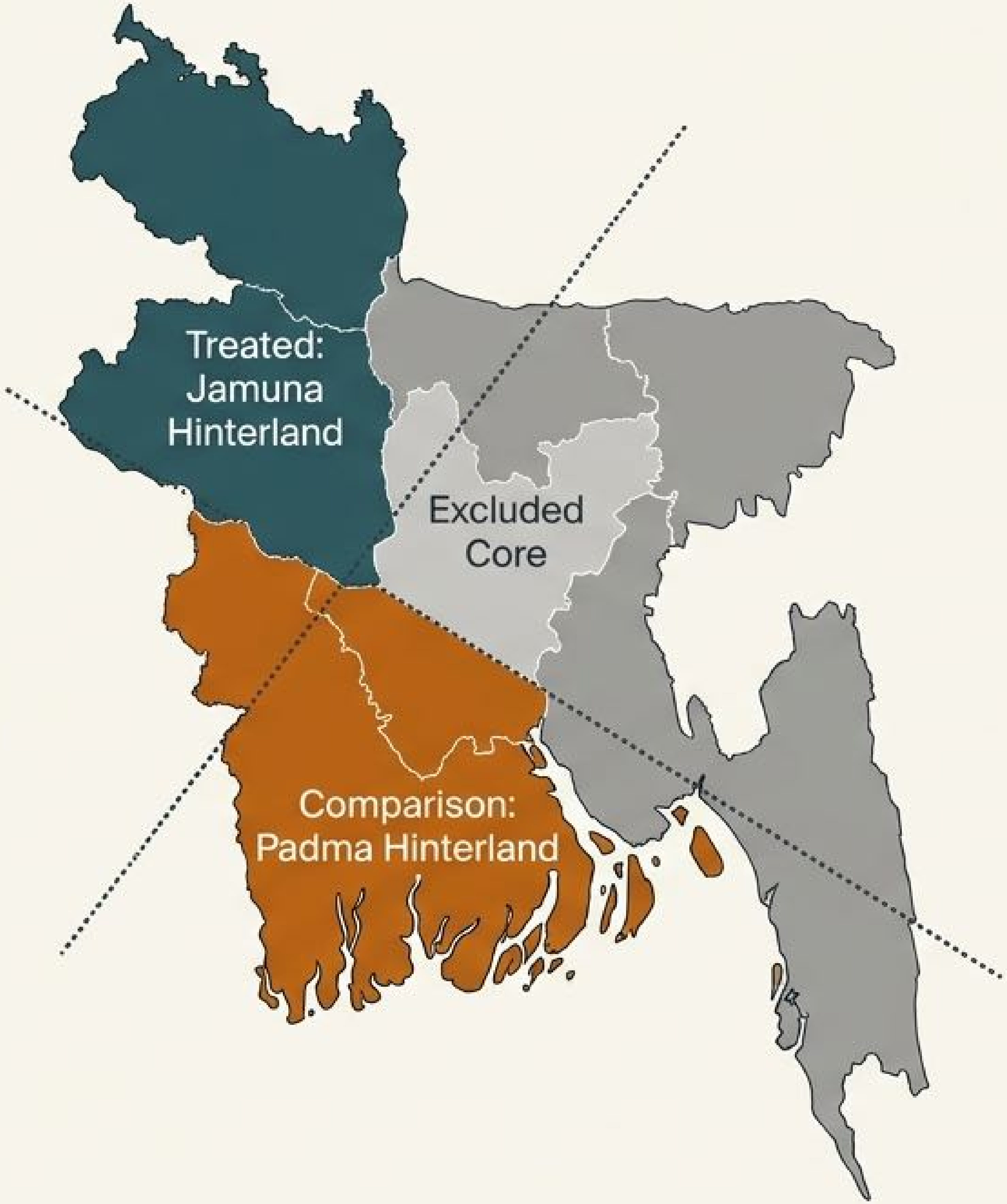
Evaluating the Impact of Infrastructure

A Methodological Study Guide to Difference-in-Differences, Event Studies, and Doubly Robust Estimation

Based on the empirical evaluation of the 1998 Jamuna Bridge (Bangladesh).





The Economic Question & The Identification Strategy

The Shock:
In June 1998, a 4.8km bridge opens over the Jamuna river, cutting freight time to Dhaka from 20 hours to 6 hours for 26 million people.

The Comparison Group:
The Padma Hinterland. Perfectly symmetric, isolated by a similar river with no bridge until 2015, and sharing the exact same agro-climatic zone.

The Identification Argument:
Bridge priority was political (driven by home-region favoritism of politicians), not economic. It did not track local pre-existing economic shocks, preventing endogenous placement bias.

The Target Estimand: Defining the ATT

Rigorous Definition

$$\tau_{ATT} = E\Big[\, Y_{it}(1) - Y_{it}(0) \mid D_J = 1, t > 1998 \,\Big]$$

Observed luminosity with the bridge.

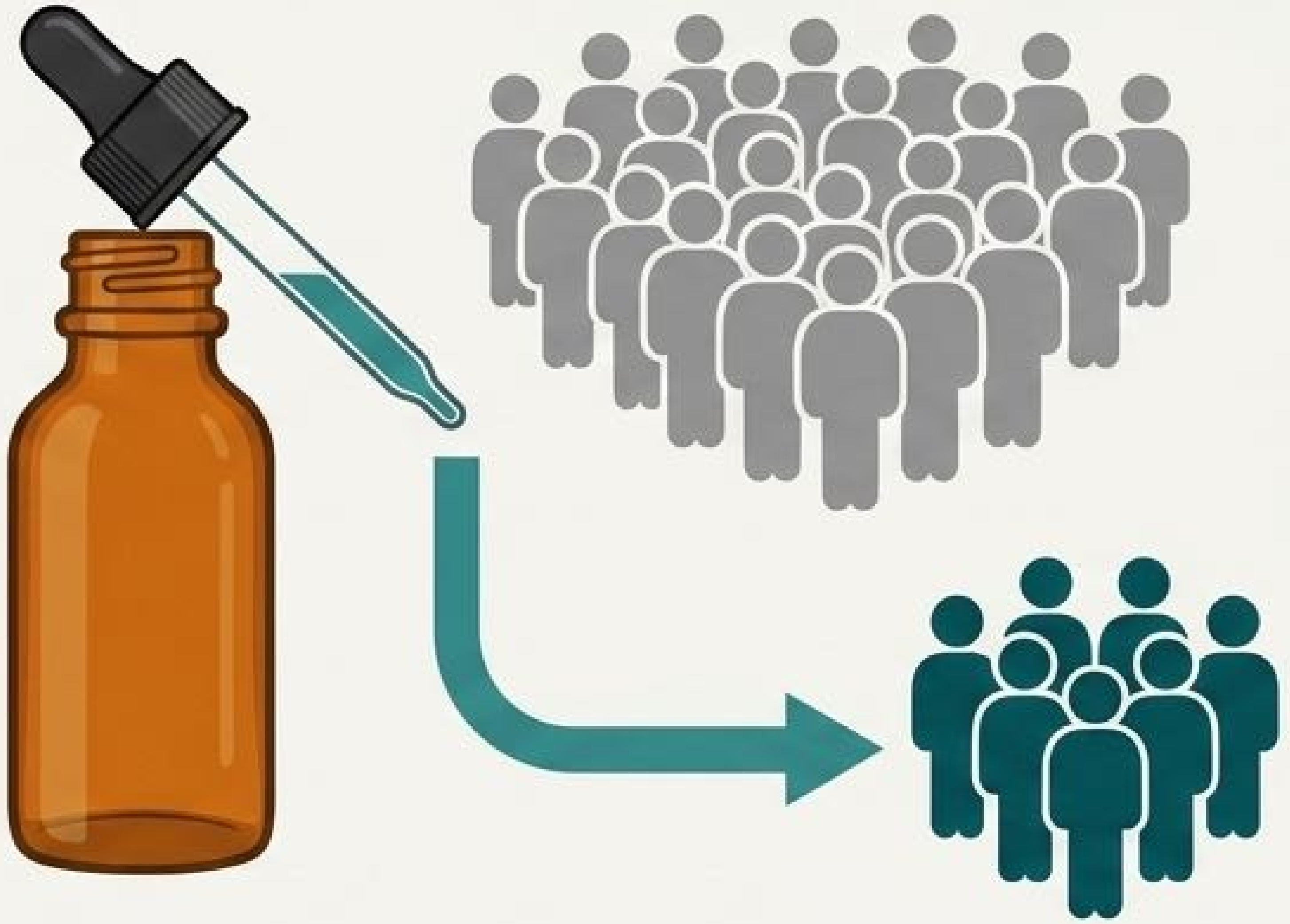
Unobserved counterfactual without the bridge.

The 123 treated upazilas in the Jamuna hinterland.

Average Treatment Effect on the Treated (ATT).

Interpretation Rule: The ATT asks “Was this specific bridge worth it?” It is not the ATE. We are not estimating what a bridge would do to a randomly chosen region in Bangladesh.

Intuition & Analogy



Analogy: Asking how much a medicine helped the specific patients who actually took it, not how it would affect the general population.

The Core Assumption: Parallel Trends & The 2x2 Logic

The Mechanism (Two Subtractions)

$$\tau_{2x2} = (Y_{J,post} - Y_{J,pre}) - (Y_{P,post} - Y_{P,pre})$$

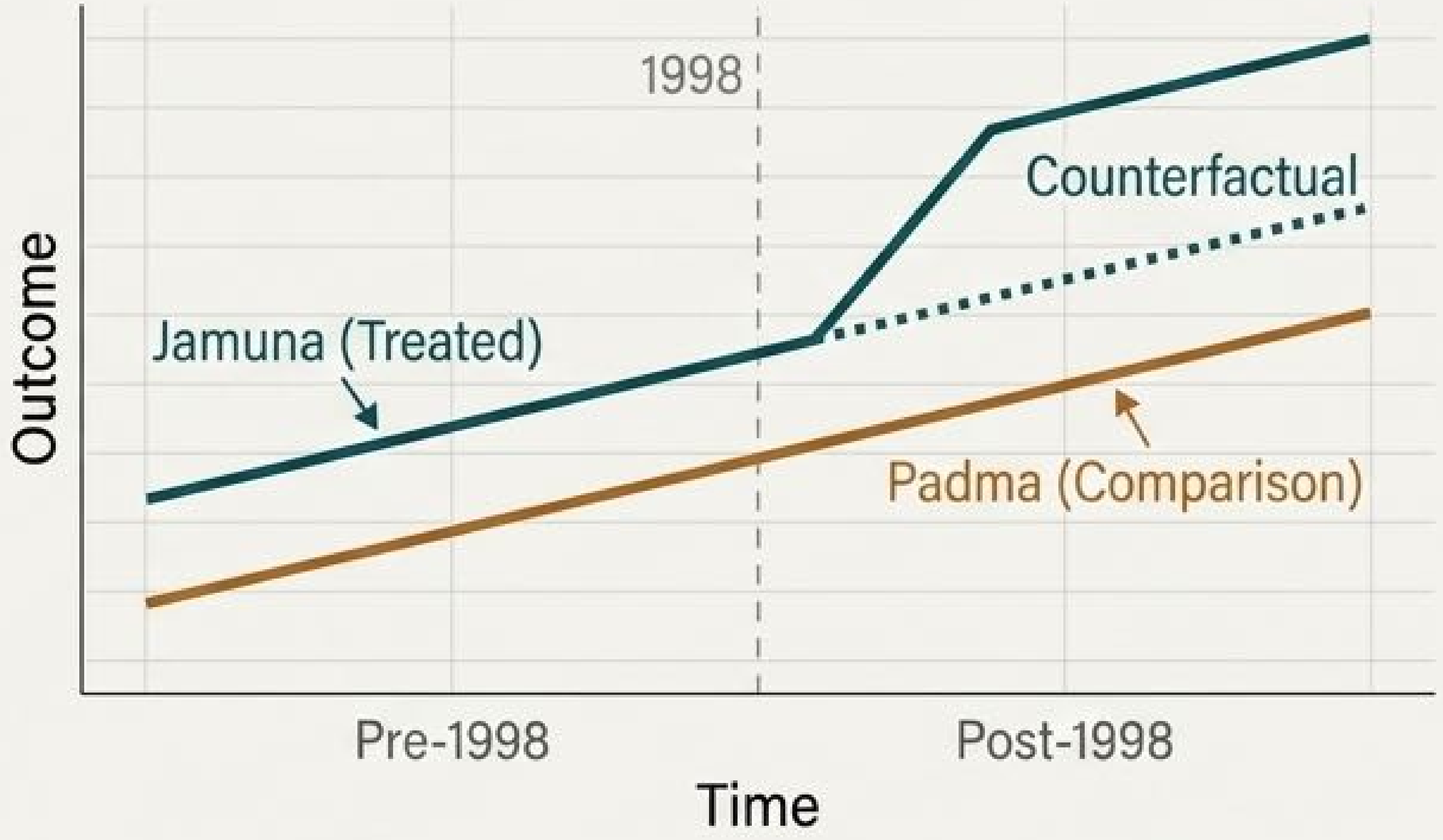
Compare how much the treated group changed against how much the untreated group changed. Anything that moved both groups equally (e.g., national shocks) cancels out entirely.

The Untestable Assumption:

$$E\left[Y_{it}(0) - Y_{i,t-1}(0) \mid D_J = 1\right] = E\left[Y_{it}(0) - Y_{i,t-1}(0) \mid D_J = 0\right]$$

Absent the treatment, the treated and comparison groups would have moved by the exact same amount.

Analogy: Two hikers on parallel ridges.



Analogy: Two hikers on parallel ridges. The permanent height gap between them cancels out; we only measure the divergence in slope caused by a helicopter lifting one of them. Parallel trends permits groups to start at different levels—they only need to move at the same rate.

Baseline Implementation: Two-Way Fixed Effects (TWFE)

The TWFE Specification:

$$Y_{it} = \theta_0 + \underbrace{\mu_i}_{} + \underbrace{\mu_t}_{} + \underbrace{\theta_1}_{}(D_J \times D_{post}) + \sum \beta_q X_{qit} + \sum \pi_m \underbrace{(Z_{mi0} \times t)}_{} + \varepsilon_{it}$$

μ_i: Upazila Fixed Effects (Absorbs permanent baseline remoteness).

θ_1: The ATT (The coefficient on the interaction term).

$Z_{mi0} \times t$: Initial conditions on a trend. Allows upazilas with different 1991 baselines to be on distinct trajectories.

μ_t : Period Fixed Effects (Absorbs national time shocks).

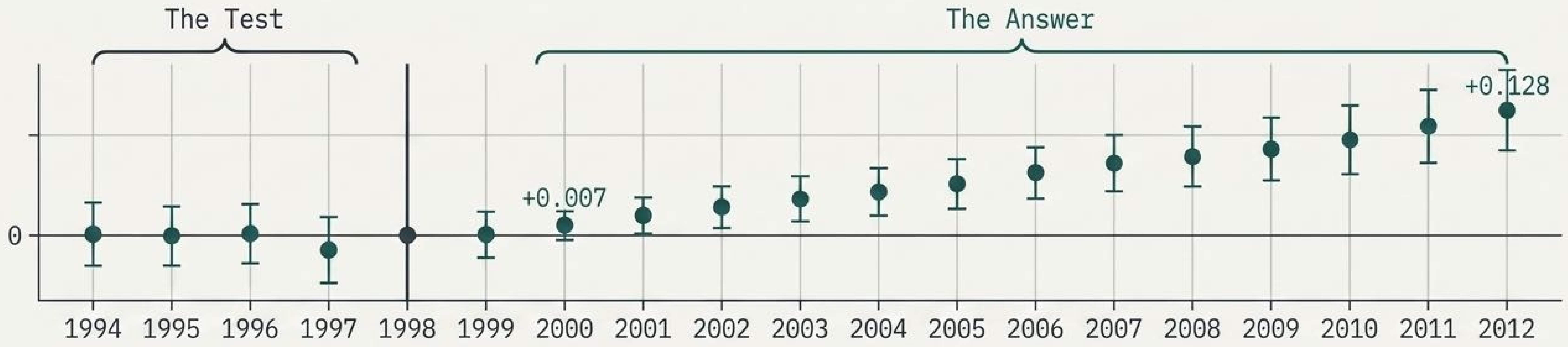
```
diff_diff(data, absorb=['geocode', 'year'], treat='treat', time='post')
```

Callout & Analogy

Implementation Warning: Do not include treatment or post dummies independently; the unit and period fixed effects mathematically absorb them.

Analogy: Grading on two curves. Compare a student to their own past average, then to the class average for that specific exam.
What survives both subtractions is the pure “textbook effect”.

Unpacking Dynamics: Event Studies



Unpacking Dynamics

Instead of one post-treatment average, estimate a separate effect for every period relative to the final pre-bridge baseline.

1. The Test (Pre-treatment)

Validates parallel trends. Coefficients should be statistically indistinguishable from zero (e.g., Nightlights pre-trend: -0.008, p=0.93).

2. The Answer (Post-treatment)

Traces the structural ramp of the effect. A single pooled average hides migration and supply-chain lags. The data shows a monotonic, 15-year climb.

Analogy

Taking a patient's temperature every single day to verify it was already stable before taking the pill, rather than just asking "are you better now?"

Correcting Imbalances: The Motivation for Doubly Robust

The Problem

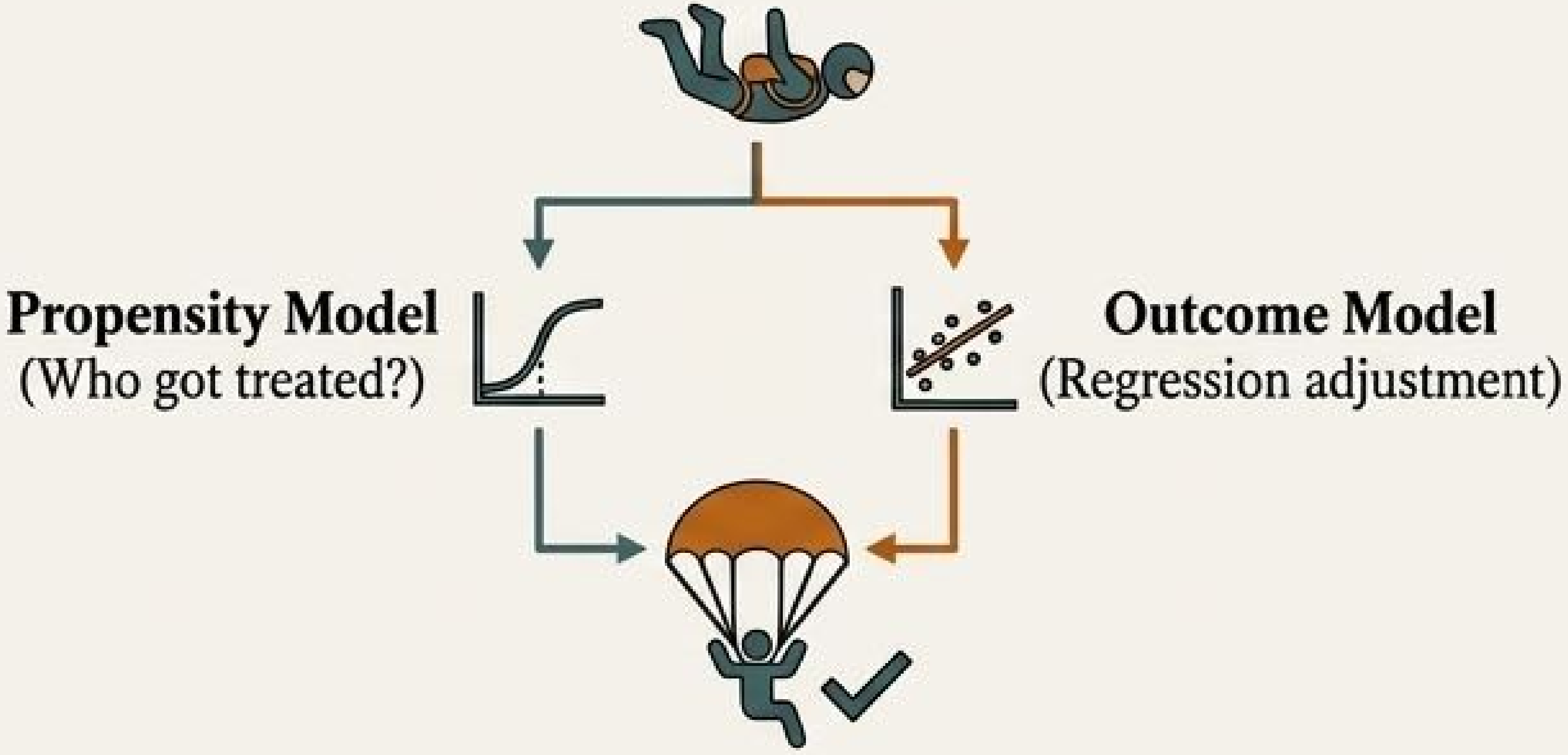
Balance Table

Baseline Characteristic	Padma (Control)	Jamuna (Treated)	
Mean Income (1990)	55.0%	33.0%	
Literacy Rate (1990)	93.7%	94.0%	
Distance to Market (km)	~1.2%	~1.3%	
Distance to Market (km)	~0.3%	~0.3%	
Distance to Market (km)	~0.3%	~0.3%	
1991 Gap	**Agriculture: +8.8%**	**Services: -8.8%**	**~0.0%**
Distance to Market (km)	~0.7%	~0.5%	
Literacy Rate (1990)	84.0%	88.0%	
Distance to Market (km)	~1.7%	~1.0%	
Literacy Rate (1990)	59.7%	46.7%	

The Motivation: The two hinterlands are perfectly symmetric geographically, but not identical economically. Remote upazilas have fundamentally different baseline trajectories. We must condition on initial characteristics.

The Solution

The Doubly Robust Parachute Model



The Doubly Robust (DR) Concept: An estimator that combines two distinct models to adjust for imbalances.

The Guarantee: The estimator is mathematically consistent if either the propensity model OR the outcome model is correctly specified.

Analogy: A skydiver jumping with both a main parachute and a reserve. Only if both models fail does the statistical jump end badly.

Doubly Robust Implementation I: Propensity-Odds & Trimming

Rigorous Logic

Step 1: The Propensity Model.

Estimate treatment likelihood based purely on pre-treatment fixed data (1991 population and distance to bridge foot).

Step 2: The 5% Trim.

Discard comparison upazilas with $p < q_{0.05}(p)$.

Why Trim?

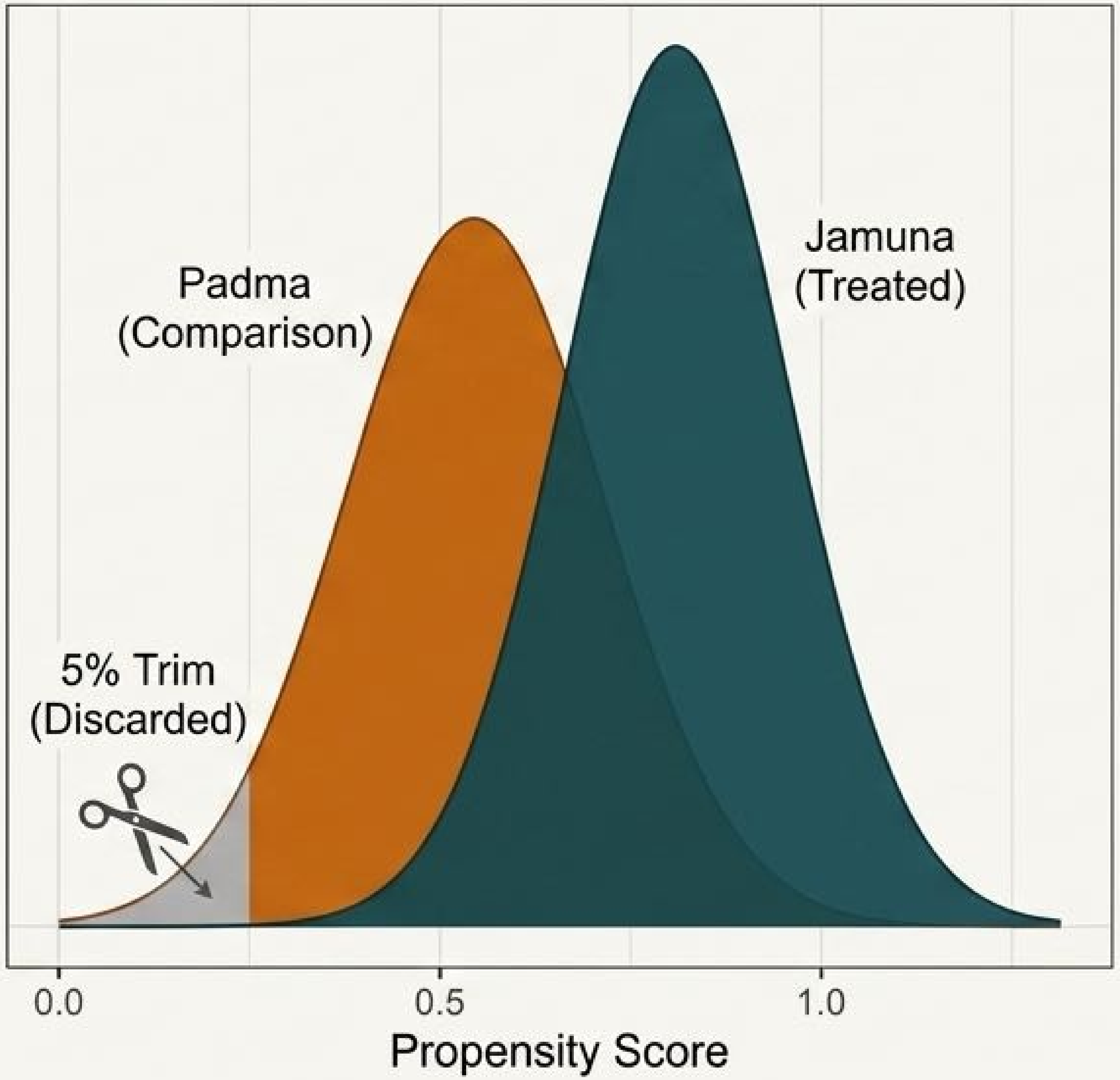
A control unit with a 0.03 propensity score must be multiplied 30x to simulate a treated unit. This amplifies enumerator errors thirtyfold. Trimming is refusing to trade at extreme statistical exchange rates.

Step 3: Propensity-Odds Weighting (LWDR).

$$w_i^{LW} = \left[\frac{p_i}{1-p_i}\right] \times \left[\frac{1-\pi}{\pi}\right].$$

Targets the ATT by weighting treated units at exactly 1.0 and reshaping the control distribution to perfectly match it.

Overlap Density Plot



Doubly Robust Implementation II: Oaxaca-Blinder (KOBDR)

Definition:

Kline's Oaxaca-Blinder Reweighting (KOBDR) mathematically projects the comparison group's outcome model onto the exact covariate profile of the treated group.

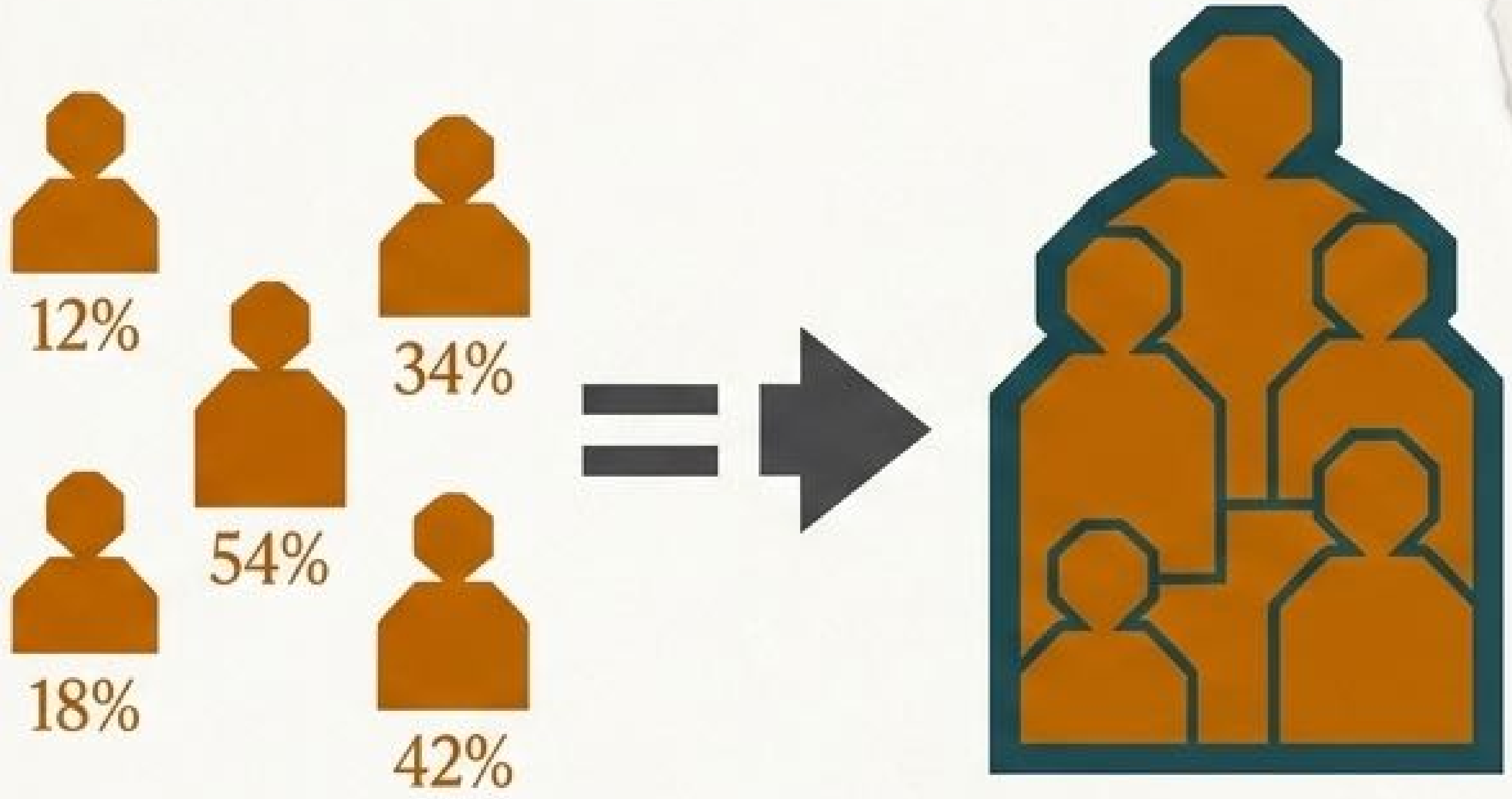
The Two-Step Procedure:

1. Run outcome regression on comparison units only.
2. Evaluate it at the average treated covariate profile.

Implementation Note:

The math can sometimes produce negative weights ("negative casting"). These units must be dropped before the final regression.

Result: Both LWDR and KOBDR converge precisely: adjusting for baseline differences raises the nightlights estimate from an unweighted 0.088 to 0.109.



Analogy: Casting understudies.

A director must stage the Jamuna region but is only allowed to hire Padma actors. They calculate what precise blend of Padma actors exactly reproduces the Jamuna demographic profile.

Interpreting the Mean Results: Reading the Math

Variable	Coefficient
Nightlights (Logged)	+0.109
Manufacturing Share	-0.010

Log Transformations (Nightlights)

Because the outcome is logged, a coefficient of 0.109 must be exponentiated.

Percentage Change = $100(e^{0.109} - 1) = 11.5\%$ increase.

Shares vs. Percents (Employment)

Employment shares are not logged. A coefficient of -0.010 is an absolute 1.0 percentage point drop.

Contextualizing the Base: A 1.0 point fall in manufacturing sounds trivial until contextualized against the 1991 baseline of only 2.8%. The bridge wiped out one-third of the entire manufacturing sector in the hinterland. Deindustrialization is massive and real.

Synthesis: The Discriminating Test (Theory vs. Data)

	Industrial Share	Population Density
Big Push	↑	↑
Backwash	↓	↓
Comparative Advantage	↓	↑

THE DATA RESULT: Manufacturing fell 1.2 points (Down), but Population Density rose 5.9% (Up).

The Theoretical Trap: Both Backwash (decay) and Comparative Advantage (specialization) predict that manufacturing will leave the periphery. Looking at factories alone yields a false conclusion.

The Deciding Variable: Population Density. Backwash requires the periphery to empty out. Comparative Advantage requires it to grow.

The Conclusion: The Jamuna hinterland did not decay into a hollowed-out backwater. It actively specialized in its comparative advantage: agricultural processing and transport services.

The Danger of Averages: Spatial Heterogeneity

Distance Gradient Matrix

	Nearest (<84km)	Middle (84-128km)	Farthest (>128km)
Services	-2.6	+1.7	+5.9
Agriculture	+3.2	+0.8	-5.7
Rice Yields	+4.9%	+6.5%	+26.5%

The Phenomenon

Average treatment effects hide profound reversals. Nearest upazilas lost services and gained agriculture. Farthest upazilas gained services, lost agriculture, and saw yields spike 26.5%.

The Economic Logic

A 17% transport cost cut on a high-barrier, long-distance trip yields massive absolute dollar savings compared to a 40% cut on a short, cheap trip. Trade responds to the absolute levels of barriers.

Policy Implication

The highest payoff of an infrastructure megaproject often occurs at the end of the line, not adjacent to the concrete build site.

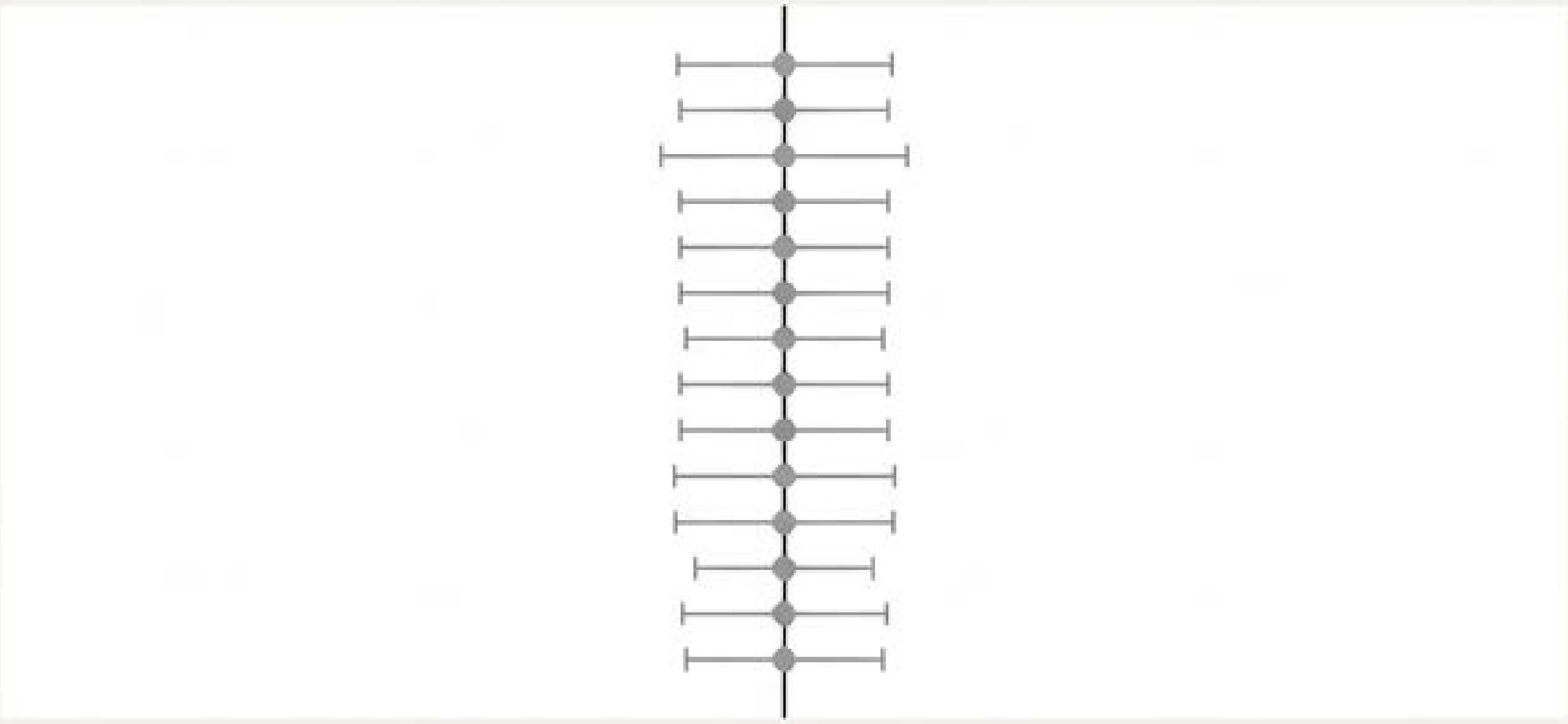
Sensitivity & Stress Testing

HonestDiD Bounds

HonestDiD (Relative Magnitude): Instead of just accepting a pre-trend test, we bound the failure. The result survives until $M \approx 1$.

Interpretation: A post-bridge unobserved shock would have to be as large as the absolute largest pre-bridge violation ever recorded to snap the causal conclusion. How many strands can snap before the bridge falls?

Public-Goods Placebo



Public-Goods Placebo: Was the "bridge effect" just a disguise for a politician flooding their home region with public money?

- **The Test:** Run the exact DiD model on village electricity access and distances to clinics/schools.
- **Result:** Null across 21 indicators.
- **Lesson:** Name the alternative mechanism and look for its specific fingerprints.

Methodological Rigor: The Replication Imperative

The Hidden Bug

```
use replication_data, clear
[MISSING: global trimL 5]
gen cut11 = r(p$trimL)
```

In the original published replication package, a missing macro silently turned the entire comparison group to missing values.

The Catastrophe

```
ATT Coefficient:        1.064 ***
Observations:             868
Number of Upazilas (N: 124
```

The regression did not throw an error. It cheerfully ran a within-group contrast on the treated units only and output a massive, incorrect coefficient.

The Tell: The coefficient was just a number. The actual warning sign was in the regression footer: N=124 upazilas instead of the full 239.

Study Guide Rule #1: Always, always read the sample size first in any regression output.

(Note: ln(0) evaluates as missing in Stata, but -inf in NumPy. Handle your nulls or corrupt your sample).

The Causal Inference Blueprint

1 Target the Right Estimand

Explicitly define whether you are estimating the ATT (this specific bridge) or the ATE (a theoretical bridge). Reweight accordingly.

2 Check Assumptions, Don't Just State Them

Plot event studies. Use HonestDiD bounds to explicitly calculate how much parallel trend failure your result can actually survive.

3 Pack Two Parachutes

Always implement Doubly Robust estimators (like KOBDR) to protect your results against isolated modeling errors.

4 Look Beyond the Average

Split your data across time (short-run vs. long-run) and space (distance bands) to unearth hidden sign reversals.

4 Look Beyond the Average

Split your data across time (short-run vs. long-run) and space (distance bands) to unearth hidden sign reversals.

5 Find the Discriminating Variable

When two theories predict the exact same primary outcome (e.g., loss of manufacturing), find the secondary variable (population density) where they violently diverge.