Regression Discontinuity in Time

Devesh Roy

International Food Policy Research Institute

ICAR-National Institute of Agricultural Economics and Policy Research (NIAP)-Winter School Program on "Advanced Tools for Impact Assessment of Agricultural Technologies and Policy Analysis"

New Delhi | March 17, 2026

The ATMI-II Project 2025-28

HISTORY OF REGRESSION DISCONTINUITY (YAMAMATO 2016)

- RDD: A fairly old idea (Thistlethwaite and Campbell, 1960)
- Recently experienced a renaissance because of formal theoretical foundation given by causal inference framework
- Applicable when treatment is assigned according to a rule based on another variable (called the forcing or running variable)
- Often useful for analysis in a “rule-based” world (administrative programs, elections)
- High internal validity: One of the few observational designs that reproduced an experimental benchmark (Cook and Wong 2008)
- Limited external validity: Effect is only identified for a small subpopulation

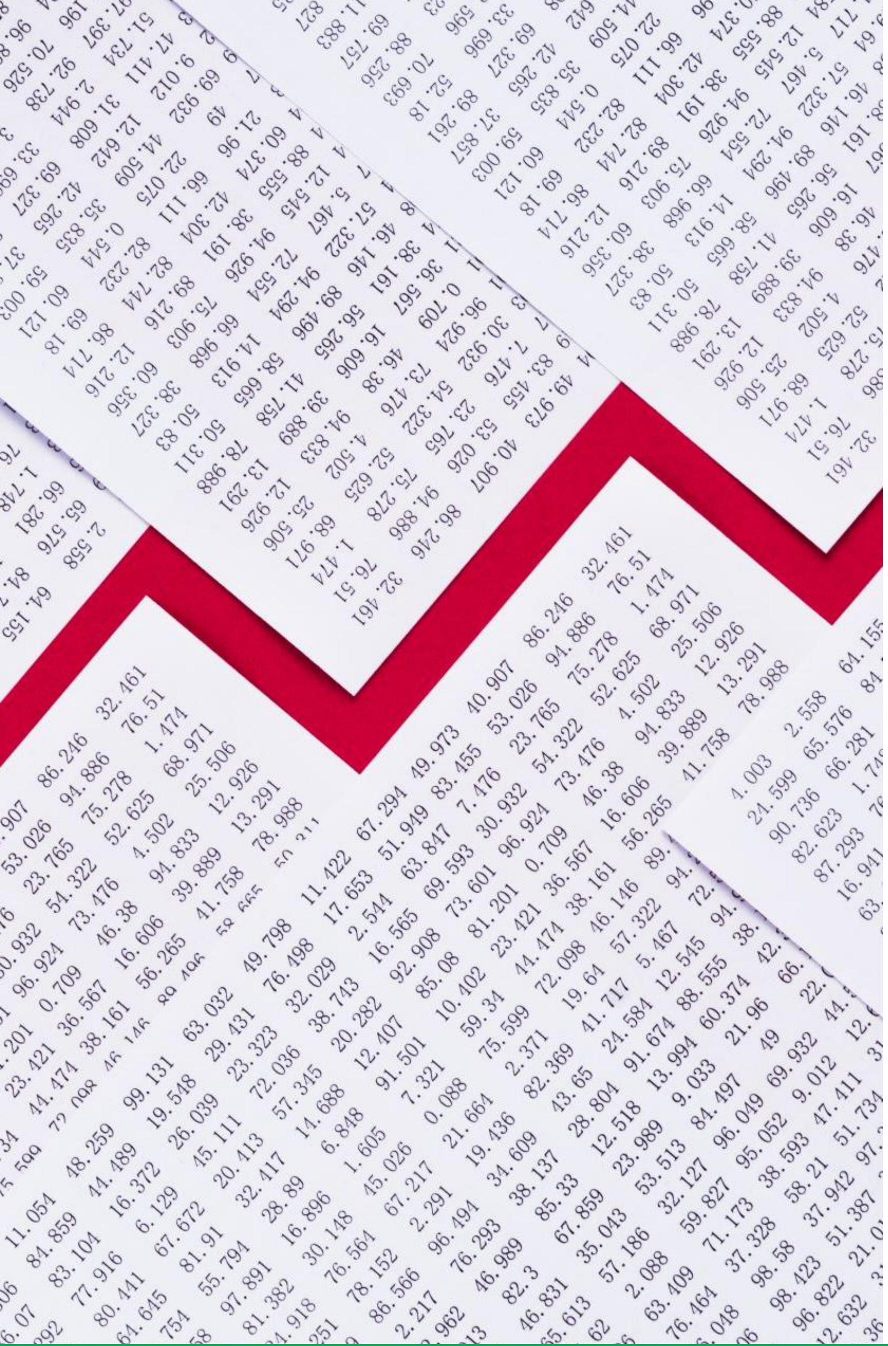
REGRESSION DISCONTINUITY DESIGN

RD Design and Applications

- Method to estimate treatment effects in natural setting
- Observed continuous variable and causal variable of interest exhibit a discontinuous increase at a certain threshold
- Address confounding factors influencing control and treatment
- Empirically verify assumptions
  - strengthens internal validity
- Applies to observations only near “cutoff point”
  - limits external validity

REGRESSION DISCONTINUITY: WORLD BANK (2014)

- When to use this method?
  - The beneficiaries/non-beneficiaries can be ordered along a quantifiable dimension.
  - This dimension can be used to compute a well-defined index or parameter.
  - The index/parameter has a cut-off point for eligibility.
  - The index value is what drives the assignment of a potential beneficiary to the treatment (or to non-treatment.)
- Intuitive explanation of the method:
  - The potential beneficiaries (units) just above the cut-off point are very similar to potential beneficiaries just below the cut-off.
  - Compare outcomes for units just above and below cutoff.



ELEMENTS OF RDD: CONTINUED

- Index must have a clearly defined cutoff score: i.e. a point on the index above or below which the population is classified as eligible for the program. For example, households with a poverty index score of less than 50 out of 100 might be classified as poor.
- The cutoff must be unique to the program of interest i.e. there should be no other programs, apart from the program to be evaluated, that uses the same cutoff score.
    - If a poverty score below 50 qualifies a household for a cash transfer, health insurance, and free public transportation, we would not be able to use the RDD method to estimate the impact of the cash transfer program by itself.
- The score of a particular individual or unit cannot be manipulated by enumerators, potential beneficiaries, program administrators, or politicians.
- It is an experiment around the cutoff

THE CUTOFF RULE

- All persons on one side of the cutoff are assigned to one group.
- All persons on the other side of the cutoff are assigned to the other
- Need a continuous quantitative preprogram measure.
- Social programs often use an index to decide who is eligible to enroll in the program and who is not. For example, antipoverty programs are typically targeted to poor households, which are identified by a poverty score or index.
- Test scores are another example. College admission might be granted to the top performers on a standardized test, whose results are ranked from the lowest to the highest performer.
- In both examples, there is a continuous eligibility index (poverty score and test score, respectively) that allows for ranking the population of interest, as well as a threshold or cutoff score that determines who is eligible and who is not.

EXAMPLE: EFFECT OF CASH TRANSFER ON CONSUMPTION (WORLD BANK 2014)

Goal

Target transfer to poorest households

Method

- Construct poverty index from 1 to 100 with pre-intervention characteristics
- Households with a score ≤ 50 are poor
- Households with a score >50 are non-poor

Implementation

Cash transfer to poor households

Evaluation

Measure outcomes (i.e. consumption, school attendance rates) before and after transfer, comparing households just above and below the cut-off point.

IDENTIFICATION FOR SHARP DISCONTINUITY

$$y_i = \beta_0 + \beta_1 D_i + \delta(score_i) + \varepsilon_i$$

$$D_i = \begin{cases} 1 & \text{If household } i \text{ receives transfer} \\ 0 & \text{If household } i \textit{ does not} \text{ receive transfer} \end{cases}$$

$\delta(score_i) =$ Function that is continuous around the cut-off point

- Assignment rule under sharp discontinuity:

$D_i = 1 \longleftrightarrow score_i \leq 50$

$D_i = 0 \longleftrightarrow score_i > 50$

IDENTIFICATION FOR FUZZY DISCONTINUITY RULES DETERMINE ELIGIBILITY BUT THERE IS A MARGIN OF ADMINISTRATIVE ERROR.

$$y_i = \beta_0 + \beta_1 D_i + \delta(score_i) + \varepsilon_i$$

$$D_i = \begin{cases} 1 & \text{If household receives transfer} \\ 0 & \text{If household } \textit{does not} \text{ receive transfer} \end{cases}$$

- But
  Treatment depends on whether $score_i >$ or $<$ 50

- And
  Endogenous factors

IDENTIFICATION FOR FUZZY DISCONTINUITY

$$y_i = \beta_0 + \beta_1 D_i + \delta(score_i) + \varepsilon_i$$

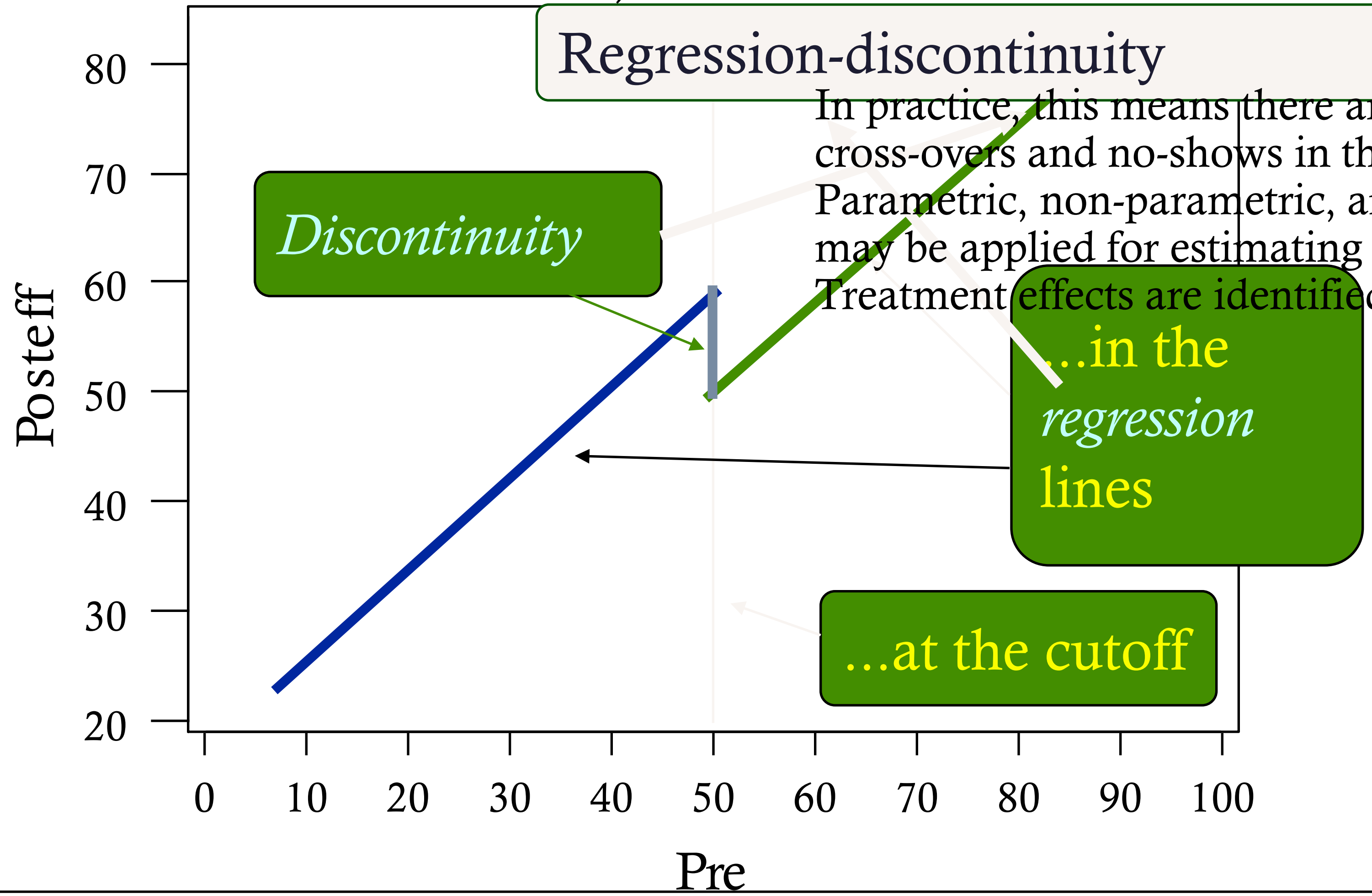
IV estimation

- First stage: $$D_i = \gamma_0 + \gamma_1 \underbrace{I(score_i > 50)}_{\text{Dummy variable}} + \eta_i$$

- Second stage: $$y_i = \beta_0 + \beta_1 D_i + \underbrace{\delta(score_i)}_{\text{Continuous function}} + \varepsilon_i$$

If there is a treatment effect, there will be a…

Regression-discontinuity

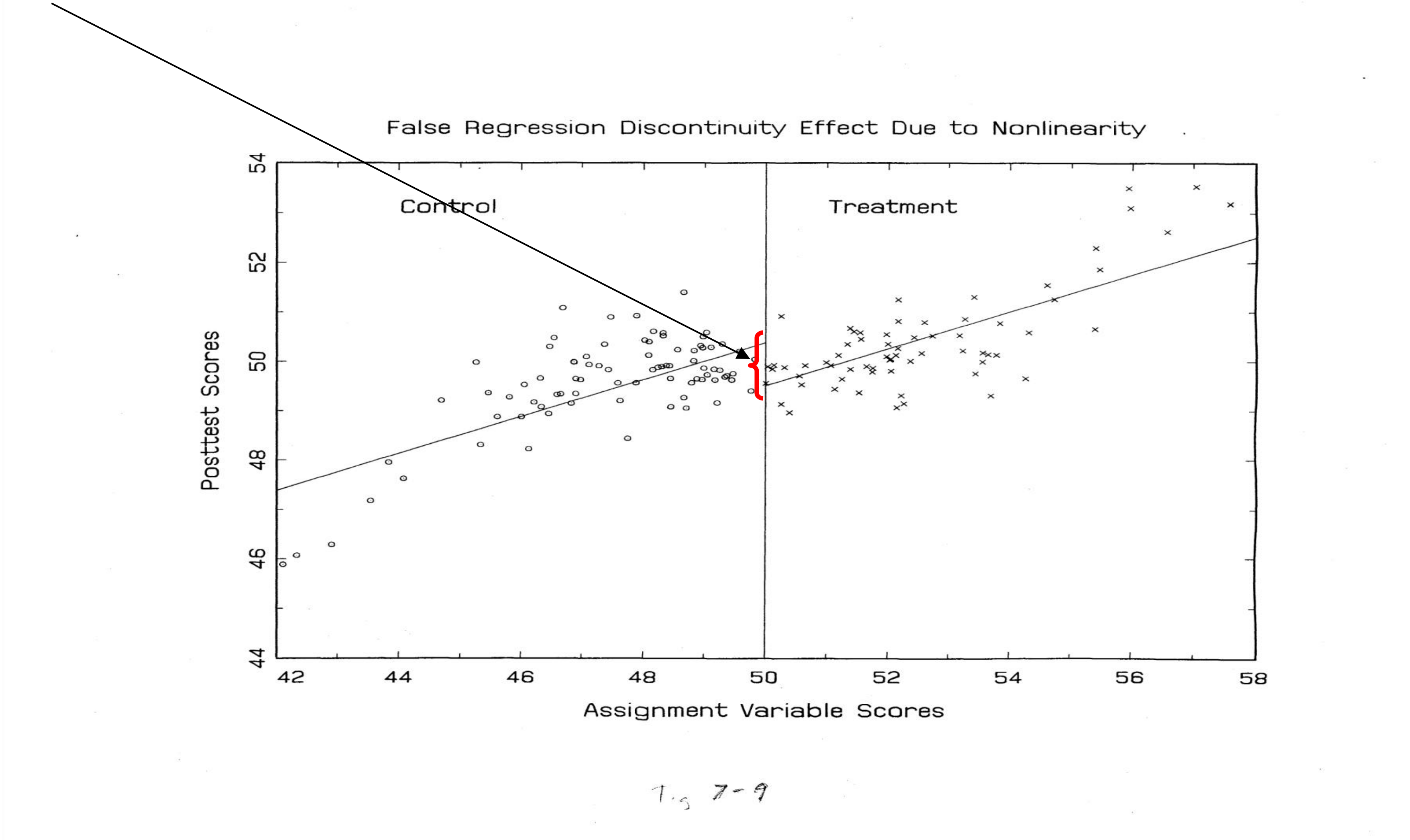


In practice, this means there are no cross-overs and no-shows in the study.
Parametric, non-parametric, and semi-parametric approac
may be applied for estimating treatment effects.
Treatment effects are identified at the cutoff.

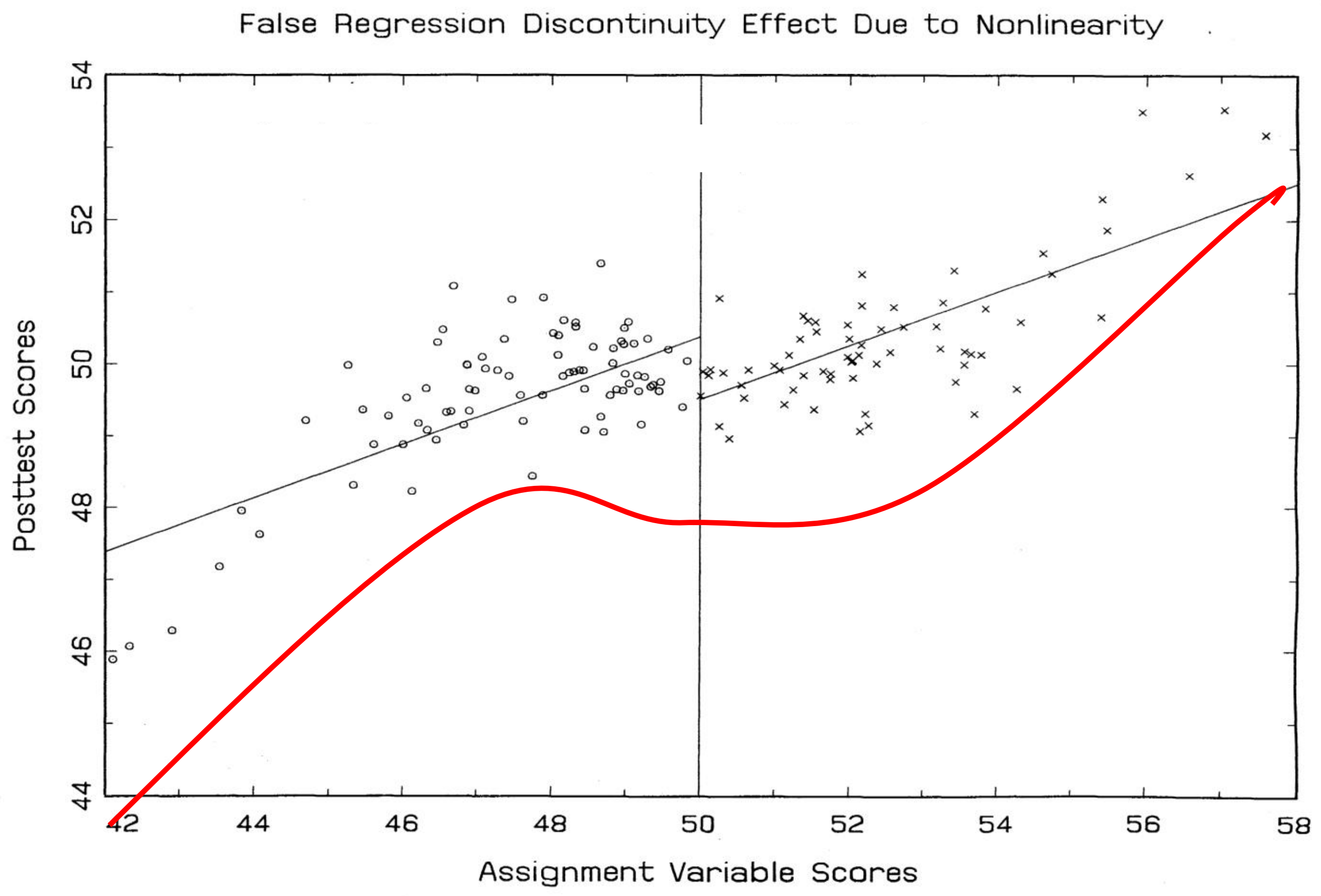
FUNCTIONAL FORM: INTERACTIONS

- Sometimes the treatment works better for some people than for others
  - For example, it is common to find that more advantaged children (higher SES, higher pretest achievement scores) benefit more from treatment than do less advantaged children.
- If this interaction (between the assignment variable and treatment) is not modeled correctly, a false discontinuity will appear:
- Anything that affects the size of that discontinuity other than treatment is a threat.
- In the example, we assumed the relationship between assignment and outcome was linear—regressions are straight lines.
- But functional form can be nonlinear due to:
  - Nonlinear Relationships between the assignment variable and the outcome
  - Interactions between the assignment variable and treatment.
- Put more technically, effects are unbiased only if the functional form of the relationship between the assignment variable and the outcome variable is correctly modeled
- Consider first the example of a nonlinear relationship between the assignment variable and the outcome variable:

Here we see a discontinuity between the regression lines at the cutoff, which would lead us to conclude that the treatment worked. But this conclusion would be wrong because we modeled these data with a linear model when the underlying relationship was nonlinear.



If we super-impose a nonlinear regression line[1] onto the data, a line that seems to match the curves in the data pretty well, we see no discontinuity at the cutoff anymore, and correctly conclude that the treatment had no effect.



[1] In this case, a cubic function (X^3)

WHAT TO DO?

- Conduct graphical analyses
- Examine residuals
- Check to see whether treatment effects are robust to alternative specifications of functional form
- Check to see whether treatment effects are robust to alternative pseudo-cutoffs
- Estimate treatment effects using non-parametric and semi-parametric approaches

LOCAL AVERAGE TREATMENT EFFECT IN RDD (SUPPOSE AN AGRICULTURAL PROGRAM WITH 5 HECTARES LAND CUTOFF)

Since the comparison group is made up of farms just above the eligibility threshold, the impact given by a RDD is valid only locally—that is, in the neighborhood around the eligibility cutoff score.

Thus we obtain an estimate of a local average treatment effect (LATE)

IMPORTANT ELEMENTS OF RDD (GERTLER ET AL 2012)

- The index must rank people or units in a continuous or "smooth" way.
- By contrast, variables that have discrete or "bucket" categories that have only a few possible values or cannot be ranked are not considered smooth. Examples of the latter include employment status (employed or unemployed), highest education level achieved (primary, secondary, university, or postgraduate), car ownership (yes or no), or country of birth.

MERITS DEMERITS OF RDD (GERTLER ET AL 2012)

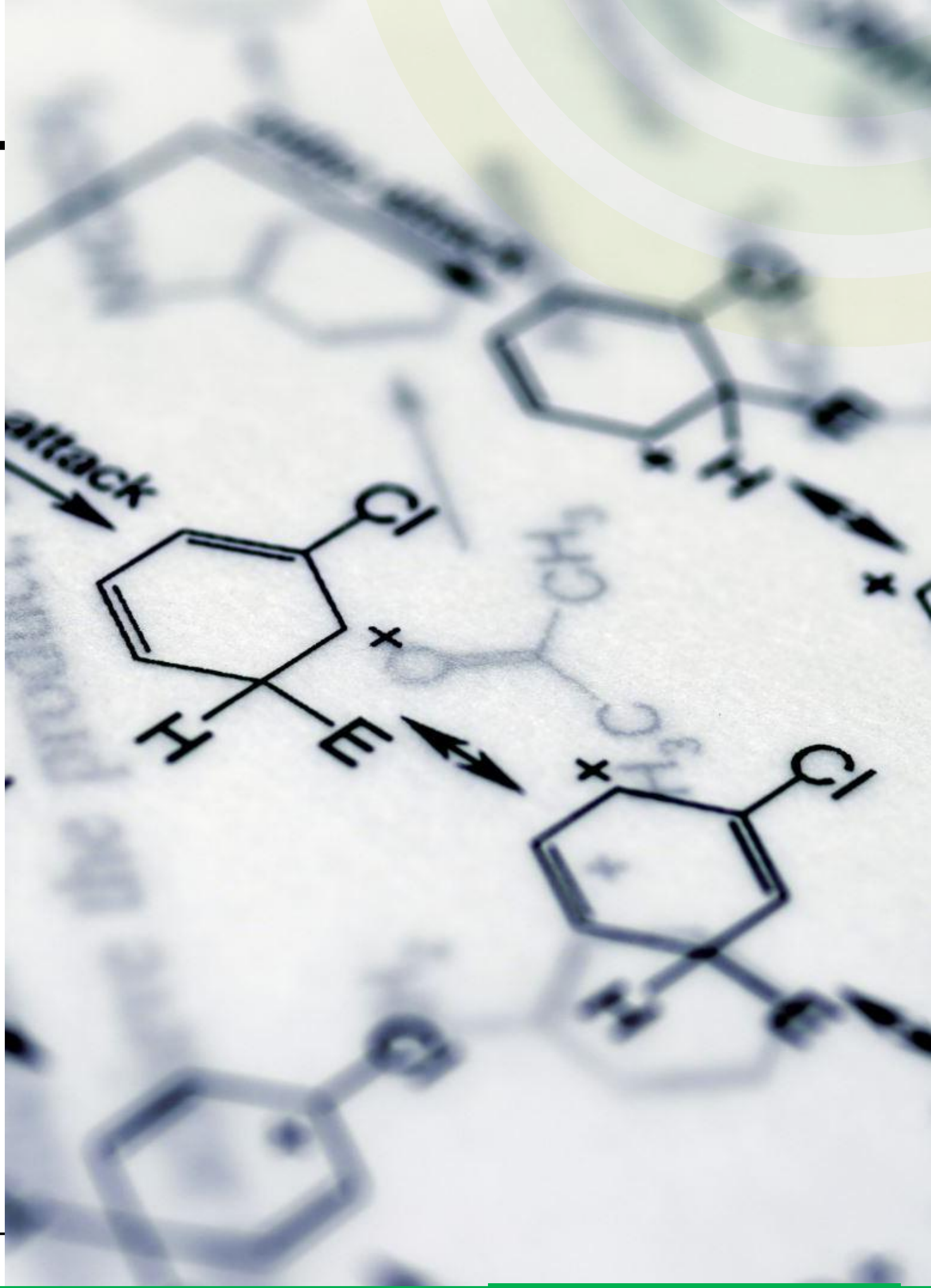
- RDD cannot provide ATE can be seen as both a strength and a limitation of the method, depending on the evaluation question of interest.
- If the evaluation primarily seeks to answer the question, should the program exist or not?, then ATE for the entire eligible population may be most relevant parameter, and clearly RDD will fall short
- However, if the policy question of interest is, should the program be cut or expanded at the margin?—that is, for (potential) beneficiaries right around the cutoff —then the RDD produces precisely the local estimate of interest to inform this important policy decision.
- Statistical power issues

DONUT RD

- **Donut RD: Excluding Observations Near the Cutoff**
- To assess whether strategic behavior or anticipation effects are influencing the estimates, researchers can use a **donut RD** approach (Barreca et al. 2011).
- This involves **removing observations immediately around the threshold** to check whether results remain consistent.
- If avoiding selection close to the cutoff significantly alters the findings, this suggests that **sorting, anticipation, or measurement error** may be affecting the estimates.
- If results are stable even after excluding these observations, it strengthens confidence in the identification strategy.

CASE FOR RDIT (DRAWN FROM TILBURG)

- Regression discontinuity in time (RDiT) designs are those RD applications where time is the running variable.
- The cutoff consequently is the treatment date: after it, subjects are treated, and before it, they are not.
- The RDiT design is useful in cases where there is no cross-sectional variation in treatment status; that is, on a certain date, treatment is applied to all subjects, meaning designs such as difference-in-differences (DiD) are not applicable.
- Like in canonical (continuity-based) RD designs, in RDiT we identify the treatment effect as a discontinuity in the outcome variable at the cutoff, assuming that any potential time-varying confounders change smoothly around the threshold.
- However, there are also important differences between standard RD and RDiT designs which have implications for causal inference.



RDIT ESTIMATES

- Estimates retrieved from RDiT are of a compound effect: the causal treatment effect of interest (i.e. what we try to retrieve from, say, a randomized controlled trial) and any unobserved sorting/anticipation/adaptation/avoidance effects that may exist but cannot be tested for.
- The extent to which the results should be interpreted solely as the causal treatment effect of interest depend on the researcher's ability to make a compelling case that the sorting effects are not present.
- Furthermore, the absence of these effects is a necessary but insufficient condition for identification

TIME SERIES ISSUES IN RDIT

- As time is the running variable, the data will have time series properties.
- This can translate into two issues:
- autoregression in the outcome variable (i.e. the outcome depends on its previous values) and serial correlation in the residuals of our fitted RD models (i.e. residuals from observations next to each other are not independent of each other, violating the Gauss-Markov assumptions).
- The first implication of this serial dependence relates to inference.
- If there is serial dependence in the residuals, standard errors must account for it. Existing RDiT literature has generally addressed this by using clustered standard errors.
- Second implication-autoregression in dependent variable (even after accounting for serial correlation in the exogenous variables and the residuals) will impact estimation of short-run versus long-run effects. Using dissipation rates short and long run effects are computed

TIME SERIES ISSUES IN RDIT

- Test for serial correlation in residuals, for instance using a Durbin-Watson test. If serial correlation is present, compute heteroskedasticity and autocorrelation consistent (HAC) standard errors, also known as Newey-West standard errors.
- In some applications it may also be relevant to test for autoregression in outcome variable and include lagged values of the outcome variable if autoregression is indeed present.
- To avoid using an excessively wide bandwidth, estimate treatment effects using Hausman and Rapson's (2018) 'augmented local linear' methodology.
- In the first stage, use the full data sample to regress the outcome on covariates/potential confounders only, excluding the treatment variable.
- In the second stage, regress the residuals of the first stage on the treatment, using an appropriate narrow bandwidth. See section 3 of Hausman & Rapson (2018) for further reference.
- Local air pollution, for instance, can dissipate in minutes or days or weeks, depending on the pollutant and local atmospheric conditions (MacDonell et al., 2014)- dynamic effects not generally included in RDIT

SORTING AND ANTICIPATION EFFECTS

- Finally, in a cross-sectional RD, a density test (e.g. McCrary (2008)) is a key check for sorting behavior.
- It is generally used to rule out selection into or out of treatment, thus making it unnecessary to further control for it.
- When time is the running variable, however, it is generally not possible to test for such behavior around the threshold.
- While the researcher can check for discontinuities in other covariates at the threshold, and for discontinuities in the outcome variable at other thresholds, the researcher cannot check for discontinuities in the conditional density of the running variable. That the density of the running variable (time) is uniform renders such tests logically irrelevant.

Table 5: Summary of Concerns for RDiT Practitioners

oncern	Intuition
nobservables correlated with time	Covariates are more important than in many cross-sectional RDs. Even with covariates included, bias is possible—for instance, a global polynomial control may overfit.
ime-varying treatment effects	Mis-specification of the treatment effect will lead to bias, particularly when using a global polynomial control.
utoregressive properties	Short- and long-run treatment effects will differ. Unless the nature of autoregression is known, estimates could approximate the short- or long-run effects, or neither.
election and strategic behavior	The running variable follows a uniform distribution across the discontinuity. It is thus impossible to test for sorting/selection around the threshold (e.g. the McCrary test).

SUMMARY OF CONCERNS ON RDIT

HOW TO DETECT NONLINEARITIES

- Visual Inspection of relationship between assignment and outcome prior to treatment (e.g., if archival data is used).
- Visual Inspection of the Graph
- Computer Programs (e.g, Cook and Weisberg)
- When in doubt, start by overfitting the model:
  - Add more polynomial and interaction terms than you think are needed, and then eliminate the nonsignificant ones (from higher order to lower order).
  - When in doubt, keep terms in the equation; such overfitting will yield unbiased estimates, but will reduce power the more nonsignificant terms are kept.

RDIT MOTIVATION

- Regression Discontinuity in Time is a special case of a Regression Discontinuity design where the "forcing variable" is time itself. At an exact cutoff time T∗, a policy or intervention is implemented. We compare observations **just before** and **just after** T∗ to estimate the causal effect.
- **Key assumptions** for RDiT:
- **Sharp assignment**: The intervention precisely begins at time T∗.
- **Local continuity**: Units just before and after T∗ are comparable except for treatment status.
- **Continuity of Time-Varying Confounders**
  - The fundamental assumption in RDiT is that unobserved factors affecting the outcome evolve smoothly over time.
  - If an unobserved confounder changes discontinuously at the cutoff date, RDiT will attribute the effect to the intervention incorrectly.
- **No other confounding interventions** that begin exactly at T∗.
- **No Manipulation of the Running Variable (Time)**
  - Unlike standard RD, where subjects may manipulate their assignment variable (e.g., test scores), time cannot be directly manipulated.
  - However, strategic anticipation of a policy (e.g., firms adjusting behavior before a tax increase) can create bias.

SOME FUNDAMENTALS FOR RDIT SETTING

- First, in RDiT it is often the case that we only have one observation per unit of time, unlike in other applications of RDD with a discrete running variable.
- This happens, for example, with hourly pollution data in environmental economics, a field where RDiT designs are frequently used-could you think of cases in our world
- In these cases, to attain sufficient sample size, researchers often use very wide bandwidths, of up to 8-10 years. In such large time periods, many changes unrelated to treatment likely to affect outcome variable.
- Time-varying confounders (in the pollution example, think of weather or fuel prices) or simply time trends in the outcome. Using a long time period after the treatment date can also result in us capturing treatment effects that vary over time, a circumstance RD designs are unsuited to.
- In addition, in RDIT time variable is measured sparsely, i.e. at low-frequency units like months or quarters, the usual assumptions of smoothness and continuity of the conditional expectation function relating the running variable to the outcome (and potential confounders) are a priori less likely to hold.
- Intuitively, almost any variable is less likely to vary smoothly from month to month than from week to week.

WHAT ARE GAUSS MARKOV ASSUMPTIONS

- The Gauss-Markov assumptions are a set of five key criteria for the Ordinary Least Squares (OLS) estimator to be the Best Linear Unbiased Estimator (BLUE).
- These assumptions are:
- 1) linear parameters,
- 2) random sampling,
- 3) no perfect collinearity,
- 4) zero conditional mean, and
- 5) homoscedasticity (constant variance)

MCCRARY TEST

- The McCrary test (2008) is a diagnostic tool for Regression Discontinuity Designs (RDD) that checks for manipulation of the running variable at the cutoff.
- It tests if the density of the running variable is continuous, as a jump (discontinuity) indicates units sorting into treatment.
- A significant result ($p<0.05$) suggests that RDD assumptions are violated, implying that the treatment assignment is not truly random near the threshold
  - Basically a statistical test for assessing whether there is a discontinuity in the density of observations at the cutoff.
  - Test is "reassuring" for the analyst, but not sufficient for proving the validity of the design.
  - Must combine thoughtful consideration of the assignment process *with* observation of distribution of data around the cutoff.

RDIT AS LOCAL RANDOMIZATION

- RDiT design not equate to local randomisation approach, cannot take time as being assigned randomly in a bandwidth around a threshold.
- Instead, we can only use the continuity-based approach to RD designs in an RDiT setting, using discontinuities in the outcome at the cutoff to identify causal treatment effects.
- Finally, as time has a uniform density, it is impossible to conduct standard density tests (such as the McCrary test) for manipulation of the running variable in RDiT designs.
- This makes it impossible to test for selection into treatment in the form of anticipation, sorting, or avoidance, all of which could bias the treatment estimates.

ASPECTS OF THE MCCRARY TEST:

Purpose: To detect if individuals or units can manipulate their assignment to treatment, which would violate the validity of an RDD.

Method: It uses local linear regressions to estimate the density of the running variable on both sides of the cutoff and tests for a statistically significant difference.

Implementation: Commonly implemented via the DC density function in Stata or dc_test in R (rdd package).

Output: The test provides a point estimate (of the log difference in heights of the density curve at the threshold, along with a standard error and z-statistic)

POTENTIAL SOURCES OF BIAS IN RDIT DESIGNS

- Extending the bandwidth-means including observations far away from treatment date, so our regressions pick up time trends and potentially time-varying treatment effects.
- Not accounting for time trends, we risk them affecting and biasing our treatment effect estimates.
- In case of time-varying treatment effects, regression to right of the cutoff (i.e. post-treatment) will be computed based on some weighted average of short- and long-run treatment effects. If these are significantly different, we will estimate a different discontinuity than if we had been able to focus on short-run effects, leading to bias in the treatment estimates.
- In addition, significant differences between short- and long-run treatment effects may cause our global polynomial to overfit.

OVERFITTING IN A POLYNOMIAL REGRESSION

Overfitting in polynomial regression occurs when a model is too complex (typically a degree > 3 or >4 causing it to map noise rather than the underlying pattern.

This occurs when a machine learning model is too complex and learns the specific details and noise of the training data instead of the general underlying patterns, resulting in poor performance on new, unseen data

MANIPULATION OF THE RUNNING VARIABLE

Manipulation of the running variable (selection into treatment): in an RDiT context, can take the form of sorting, anticipation, or avoidance.

Example, subjects anticipate the introduction of tougher regulation on installing polluting facilities and move forward their installation to just before the treatment date.

This can also happen in canonical RDDs, but in RDiT designs it represents a bigger problem because we cannot formally test for such behaviour.

Hence, any effects we estimate will be composed of both the true treatment effect and manipulation-related effects, if any.

MITIGATION-RECOMMENDATION FOR RDIT IN PRACTICE

To informally check for time-varying treatment effects, plot the residuals of a regression of the outcome on any relevant covariates (not including the treatment dummy in the regression) against time.

Use various specifications of the time trend (at least linear and quadratic).

If the residual plots with different time trends are substantially different, this may indicate the presence of time-varying effects. If the residuals appear to exhibit seasonal or cyclical patterns, consider including time/season fixed effects in your regression.

Conduct standard tests for sensitivity to polynomial order and bandwidth choice.

RDIT IN PRACTICE SUGGESTIONS

- Conduct placebo RD estimations on fake treatment dates and, if applicable, nearby geographic areas not subject to the treatment.
- Plot the relationship between time and any potential confounders to demonstrate continuity at the cutoff.
- Additionally, regress each potential confounder on the treatment to provide more formal proof of continuity. If there are discontinuities, include covariates capturing these time-varying confounders in your regression.
- Estimate a 'donut' RD, removing observations just before and after the treatment date, to minimise the threat to inference from any sorting, anticipation, or avoidance effects, which are likely to be concentrated in the period right around the treatment date.
- This will help to allay doubts about the possibility of manipulation/selection into treatment in biasing our estimates.

DONUT REGRESSION DISCONTINUITY DESIGNS

- 'Donut' regression discontinuity designs are frequently used when there is non-random 'heaping' in the data.
- Heaping refers to cases where certain values of the running variable have much higher densities than adjacent values.
- This is often the case when there is some type of manipulation of the running variable: think of the pollution example - if polluters realise that a new regulation will be enforced on day x (they *anticipate* the policy), they might want to install their polluting facilities on day $x - 1$. If many of them do so simultaneously, we will observe much higher values of installations on day $x - 1$ than on nearby dates.

RDIT SUMMARY

- We can use regression discontinuity in time (RDiT) designs when the treatment we are interested in is implemented at a certain date and is applied to all subjects, meaning we have no cross-sectional control group.
- However, we must be aware of the shortcomings and potential sources of bias inherent in RDiTs, namely using very wide bandwidths to generate sufficient sample size, serial correlation and autoregression, and manipulation.

Bias	Solution
Discontinuously time-varying confounders	Regress outcome on covariates to test for discontinuities; control for confounders in the regression if necessary
Excessively wide bandwidth: time-varying treatment effects	Plot residuals with different specifications of time trends to visualise time-varying effects; Augmented local linear procedure of Hausman & Rapson (2018)
Excessively wide bandwidth: time trends	Placebo tests on fake treatment dates, time/season fixed effects, various polynomials to model time-outcome relationship
Serial correlation in residuals	HAC/Newey-West standard errors
Autoregression in the outcome variable	Lagged values of the outcome in the regression
Manipulation	Donut RDD

EQUATION FORM OF RDIT

- $Y_t = \alpha_0 + \alpha_1 (T_t - T_*) + \tau D_t + \alpha_2 (T_t - T_*) D_t + \epsilon_t$,

MODEL SELECTION

- **Estimation and Model Selection**
- In Regression Discontinuity in Time, model selection is critical to accurately estimate the causal effect at the cutoff T∗. Unlike Interrupted Time Series, which models long-term trends before and after an intervention, RDiT relies on local comparisons around the cutoff. This means that:
- A narrow bandwidth (h) should be chosen to focus on observations just before and after T∗.
- Polynomial order selection should be guided by the Bayesian Information Criterion to avoid overfitting.
- Higher-order polynomials can introduce spurious curvature, so local linear or quadratic models are preferred.

MODEL

- Baseline local linear model
- $Y_t=\alpha_0+\alpha_1(T_t-T_*)+\tau D_t+\alpha_2(T_t-T_*)D_t+\epsilon_t, \text{for } |T_t-T_*|<h$
- where:
- Y_t is the outcome of interest at time t.
- T_t is the time forcing variable.
- T_* is the cutoff time when the policy/intervention occurs.

BASELINE LINEAR MODEL

- D_t is the treatment indicator:
- $D_t=1$ if $t \geq T*$ (post-intervention).
- $D_t=0$ if $t<T*$ (pre-intervention).
- h is the bandwidth, restricting analysis to observations close to $T*$.
- τ is the treatment effect, measuring the discontinuity at $T*$.
- $\alpha_1(T_t - T*)$ allows for a smooth time trend on both sides of the cutoff.
- $\alpha_2(T_t - T*)D_t$ captures any differential time trends post-treatment.
- This model ensures that the treatment effect is identified from the discontinuity at $T*$, rather than long-term trends.
-

QUADRATIC LOCAL MODEL (ALLOWING FOR NONLINEAR TRENDS)

- If the outcome variable exhibits curvature over time, a quadratic term can be added:
- $Y_t = \alpha_0 + \alpha_1 (T_t - T_*) + \alpha_2 (T_t - T_*)^2 + \tau D_t + \alpha_3 (T_t - T_*) D_t + \alpha_4 (T_t - T_*)^2 D_t + \epsilon_t$, for $|T_t - T_*| < h$
- $\alpha_2 (T_t - T_*)^2$ accounts for **nonlinear pre-treatment trends**.
- $\alpha_4 (T_t - T_*)^2 D_t$ allows for **nonlinear post-treatment effects**.
- This model is useful if visual inspection suggests a **curved relationship** near the cutoff.

AUGMENTED LOCAL LINEAR MODEL (ROBUST CONTROL FOR CONFOUNDERS)

- Following C. Hausman and Rapson (2018), an augmented approach helps control for omitted variables:
- **First-stage regression:** Estimate the outcome with all relevant control variables and compute the residuals.
- $Y_t=\delta_0+\sum_j \delta_j X_{jt}+\nu_t$
- where X_{jt} are observed covariates that could influence Y_t.
- **Second-stage RDiT model:** Use residuals from the first stage in the standard local linear RDiT model:
- $\hat{\nu}_t=\beta_0+\beta_1(T_t-T_*)+\tau D_t+\beta_2(T_t-T_*)D_t+\epsilon_t, \text{for } |T_t-T_*|<h$
- This approach removes variation explained by covariates before estimating the treatment effect.
- Bootstrap methods should be used to correct for first-stage estimation variance.
-

STRENGTHS OF RDIT

- One of the key advantages of **Regression Discontinuity in Time** is its ability to handle cases where standard Difference-in-Differences approaches are infeasible.
- This typically occurs when treatment implementation lacks cross-sectional variation, meaning that all units receive treatment at the same time, leaving no untreated control group for comparison.
- In such cases, RDiT provides a viable alternative by exploiting temporal discontinuities in treatment assignment.
- Notably, some studies combine both RDiT and DiD to strengthen identification and provide additional insights. For instance,
- Auffhammer and Kellogg (2011) applies these methods to examine how treatment effects vary across individuals and geographic space.

OTHER BENEFITS OF RDIT

- Beyond being an alternative to DiD, RDiT also offers advantages over simpler **pre/post comparisons**. Unlike naive before-and-after analyses, RDiT can incorporate flexible controls for time trends, reducing the risk of spurious results due to temporal confounders.
- Event study methods, particularly modern implementations, have improved significantly, allowing researchers to study treatment effects over long time horizons. However, RDiT still holds certain advantages:
- **Longer time horizons**: Unlike traditional event studies, RDiT is not restricted to short-term dynamics and can capture effects that unfold gradually over extended periods.
- **Higher-order time controls**: RDiT allows for more flexible modeling of time trends, including the use of higher-order polynomials, which may provide better approximations of underlying time dynamics.

Table 28.3: Comparison with Other Methods			
Method	Key Feature	Strengths	Weaknesses
Difference-in-Differences	Uses a control group	Accounts for time-invariant confounders	Requires parallel trends assumption
Event Studies	Models multiple time periods	Estimates dynamic treatment effects	Requires staggered interventions
Pre/Post Comparison	Simple before/after design	No control needed	Cannot separate treatment from time trends
Regression Discontinuity in Time	Uses time as the running variable	Flexible polynomial trends	Sensitive to polynomial choice, cannot model time-varying treatment

LIMITATIONS AND CHALLENGES OF RDIT: CONTINUED

- **Inapplicability of the McCrary Test**
- A key diagnostic tool in standard RD designs is the McCrary test (McCrary 2008), which checks for discontinuities in the density of the running variable to detect manipulation.
- Unfortunately, this test is not feasible in RDiT because time itself is uniformly distributed.
- This limitation makes it more challenging to rule out sorting, anticipation, or other forms of manipulation (sec-sorting-bunching-and-manipulation)) around the threshold.

POTENTIAL DISCONTINUITIES IN UNOBSERVABLES

- Even if the treatment is assigned exogenously at a specific time, **time-varying unobserved factors** can still introduce discontinuities in the dependent variable.
- If these unobservable factors coincide with the threshold, they may be mistakenly attributed to the treatment effect, leading to biased conclusions.
- RDiT does not naturally accommodate **time-varying treatment effects**, which can lead to specification issues. When choosing a time window:
- **A narrow window** improves the local approximation but may reduce statistical power.
- **A broader window** provides more data but increases the risk of bias from additional confounders.

ADJUSTMENTS/ASSUMPTIONS NEEDED TO ADDRESS CONCERNS OF TIME VARYING TREATMENT EFFECTS

- **The model is correctly specified**, meaning it includes all relevant confounders or that the polynomial approximation accurately captures time trends.
- **The treatment effect is correctly specified**, whether assumed to be smooth, constant, or varying over time.
- Additionally, these two assumptions **must not interact**. In other words, the polynomial control should not be correlated with unobserved variation in the treatment effect. If this condition fails, bias from misspecification and treatment heterogeneity can compound (C. Hausman and Rapson 2018, 544).

MEANS OF ROBUSTNESS

- **Visual Inspection: Raw Data and Residuals**
- Before applying any complex statistical adjustments, start with a simple visualization of the raw data and residuals (after removing confounders and time trends). If results are sensitive to the choice of polynomial order or local linear controls, it could signal time-varying treatment effects.
- A well-behaved RDiT should exhibit a clear and consistent discontinuity at the threshold, regardless of the specification used.
- If the discontinuity shifts or fades under different model choices, this suggests sensitivity to the polynomial approximation, potentially indicating bias.

SENSITIVITY TO POLYNOMIAL ORDER AND BANDWIDTH CHOICE

- A common concern in RDiT is overfitting due to high-order global polynomials. To diagnose this issue:
- Estimate the model with different polynomial orders and check whether results remain consistent.
- Compare global polynomial estimates with local linear specifications using different bandwidths.
- If findings remain stable across specifications, the estimates are likely robust. However, if results fluctuate significantly, this suggests potential overfitting or sensitivity to bandwidth choice.

PLACEBO TESTS

- To strengthen causal claims, conduct placebo tests by estimating the RDiT model under conditions where no treatment effect should exist. There are two primary approaches:
- Estimate the RD on a different location or population that did not receive the treatment. If a discontinuity is detected, it suggests that the estimated effect may be driven by factors other than the intervention.
- Use an alternative time threshold where no intervention took place. If the model still detects a significant effect, this implies that the discontinuity may be an artifact of the method rather than the treatment.
- If placebo tests reveal no significant discontinuities, this reinforces the credibility of the primary RDiT estimate.

DISCONTINUITY IN CONTINUOUS CONTROLS

- Another useful diagnostic is to **plot the RD discontinuity on continuous control variables** that should not be affected by the treatment.
- If these covariates exhibit a significant jump at the threshold, it raises concerns that other **time-varying confounders** may be driving the observed effect.
- Ideally, covariates should remain smooth across the threshold, confirming that the discontinuity in the outcome is **not** due to unobserved factors.

APPLICATIONS OF RDIT

- RDiT has been used extensively in **environmental economics**, **transportation policy**, and **public health**, where regulations or interventions are introduced at well-defined points in time.
- **Environmental Regulations and Air Quality**
- Several studies exploit sudden policy changes or emission regulations to estimate their impact on air pollution:
- Davis (2008), Auffhammer and Kellogg (2011), and H. Chen et al. (2018) all examine the impact of environmental regulations on air quality.
- Gallego, Montero, and Salas (2013) compares RDiT and DID estimates to assess the reliability of control groups in air pollution studies.
- By leveraging RDiT, these studies isolate immediate changes in pollution levels at policy thresholds while controlling for underlying trends.

COVID-19 LOCKDOWNS AND WELL-BEING

- Brodeur et al. (2021) employs RDiT to assess the impact of COVID-19 lockdowns on psychological well-being, economic activity, and public health outcomes.
- The sudden implementation of lockdown policies provides a sharp time-based discontinuity, making RDiT a natural method to evaluate their effects.
-

EXAMPLE –SUBWAYS STRIKES AND SLOWDOWNS (AER 2014)

- On October 14, 2003, Los Angeles County MTA workers went on strike, shutting down the entire transit system for 35 days.
- Anderson (2014) uses this abrupt halt in service to estimate the effects of transit provision on traffic congestion.
- Implement a RDIT design with the date as the running variable and October 14 as the discontinuity threshold.
- This design is similar in principle to the RD designs implemented by Davis (2008); Auffhammer and Kellogg (2011); Chen and Whalley (2012) and Bento et al. (forthcoming).

A BIT OF A STORY

- The 2003 strike was rooted in a disagreement with MTA mechanics over contributions to a health care fund.
- Mechanics had worked without a contract for over one year before striking on October 14.
- The strike's exact timing was exogenous in that it occurred on the first business day following the expiration of a 60-day court-ordered injunction on striking. MTA drivers, clerks, and supervisors honored the mechanics' picket line, shutting down the entire system (Streeter and Bernstein 2003).
- A small number of contract-operated MTA bus lines continued service, and the MTA contracted a "Red Line Special" bus service to duplicate part of the Red Line subway route.

METROLINK

- Metrolink also continued scheduled commuter rail service.
- However, these combined services carried an insignificant fraction of total MTA riders.
- Anecdotal evidence suggested that congestion increased substantially during the strike (Rubin 2003), and this was later confirmed in data analyses.
- The strike continued until November 18, at which point service was gradually resumed over the following week (Streeter, Bernstein, and Liu 2003).

BASIC ESTIMATION

$$y_{it} = \alpha + \beta strike_{it} + f(date_{it}) + \varepsilon_{it}. \quad (7)$$

BASIC ESTIMATION: CONTINUED

- In this equation *yit* is the average delay (in minutes per mile) for detector *i* during hour *t*, *strikeit* is a binary variable equal to unity when the strike is in effect and zero otherwise, and *dateit* is the date measured in days from the beginning of the strike.
- Identification in the RD model comes from assuming that the underlying, potentially endogenous relationship between ε*it* and the date is eliminated by the flexible function $f(\cdot)$.
- In particular, the relationship between ε*it* and the date must not change discontinuously on or near the date on which the strike begins. The RD is a sharp
- RD in that the running variable *dateit* completely determines *strikeit* .
- We set the RD threshold at the beginning of the strike rather than the end of the strike because service is restored gradually when the strike ends.
- There is thus no sharp change in the "treatment" when the strike ends.
- To estimate this model we follow Imbens and Lemieux (2008). With *dateit* normalized to be zero on the day the strike begins, we estimate local linear regressions of the form

LOCAL LINEAR REGRESSIONS

$$(8) \quad y_{it} = \alpha + \beta strike_{it} + \gamma_1 date_{it} + \gamma_2 date_{it} \cdot strike_{it} + \delta \mathbf{X}_{it} + \varepsilon_{it}.$$

ESTIMATOR

- In this specification, the function $f\ (date_{it})$ is specified as γ1 · $date_{it}$ + γ2 · $date_{it}$ · $strike_{it}$ (note that both $date_{it}$ and $strike_{it}$ vary only by date).
- The terms $date_{it}$ and $date_{it}\ strike_{it}$ should absorb any smooth relationship between the date and ϵ_{it} in the days surrounding the beginning of the strike.
- If the RD assumption is valid (i.e. ϵ_{it} does not change discontinuously when the strike begins)
- The stimate of β, the coefficient of interest, will be unbiased even without the controls $\mathbf{X}_{it}$.
- However, we include several variables in $\mathbf{X}_{it}$ to increase our estimates' precision. These additional controls include day-of-week indicators and detector fixed effects.
- The local linear regression estimates β at the time the strike begins, or at $date_{it} = 0$
- As a kernel estimator, it requires a kernel function and a bandwidth.

KERNEL SPECIFICATION

- We specify a uniform kernel (Imbens and Lemieux 2008) and use a bandwidth of 28 days on each side of the strike threshold in our base specification. The strike began on October 14, 2003, so the sample includes dates between September 16 and November 10 (excluding all weekends and holidays).
- Since identification strategy only attempts to estimate β at $date_{it} = 0$ (when the strike begins) no additional dates beyond the 28-day bandwidth surrounding October 14 enter the sample.
- As robustness can use varying bandwidths.
- In all cases we weight each detector by prestrike VMT. In practice this means each observation is weighted by ωi , which equals the length of highway covered by detector i multiplied by the average traffic flow across detector i in the prestrike period.

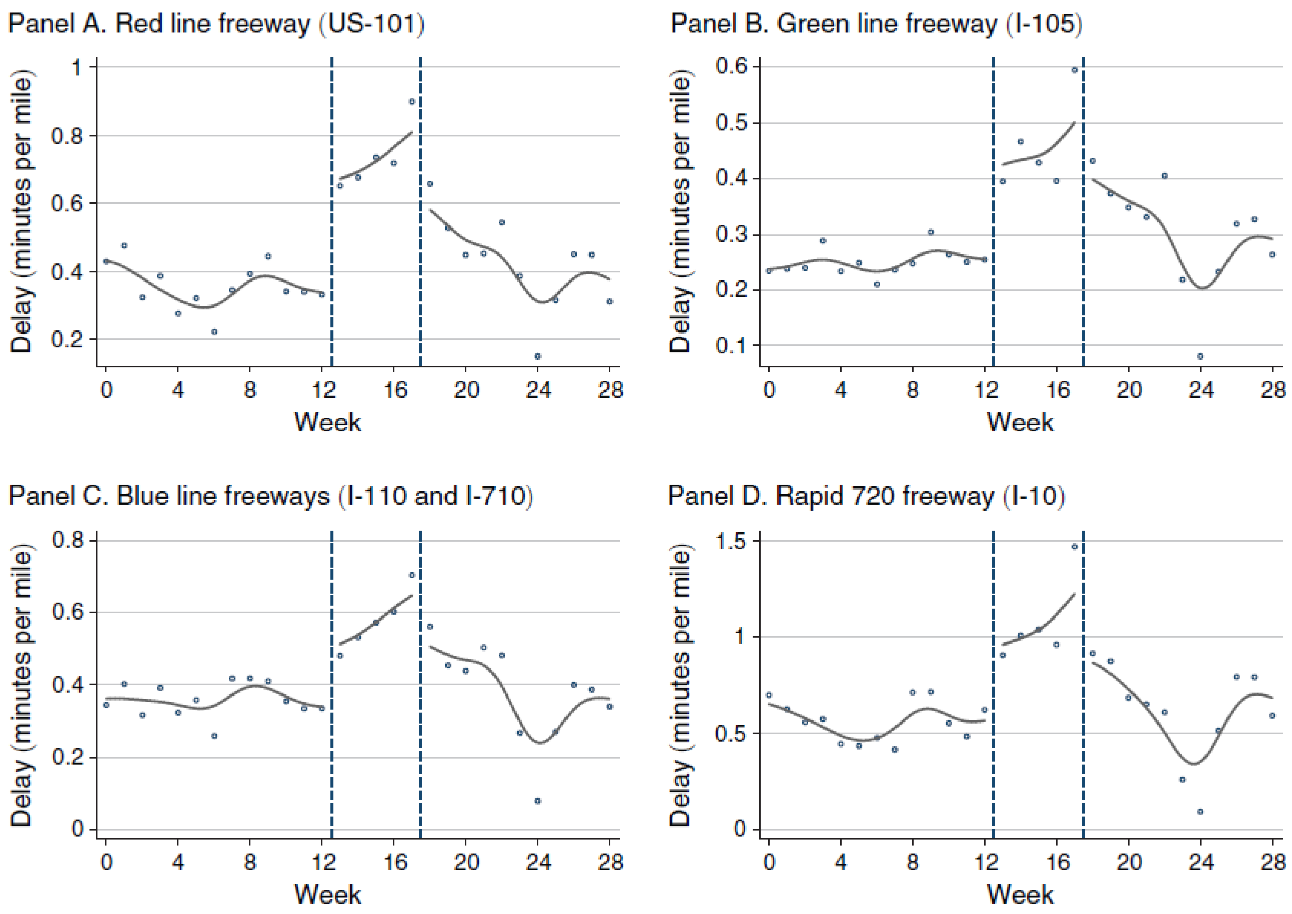


FIGURE 3. AVERAGE WEEKLY PEAK HOUR DELAY ON SPECIFIC LOS ANGELES FREEWAYS, 7/14/2003 TO 1/30/2004

RDIT RESULTS

TENETS OF IDENTIFICATION

- Identification in the RD model comes from assuming that the conditional expectation E[ε*it* | *dateit*] is smooth as *dateit* crosses the RD threshold.
- In our context this implies that factors affecting traffic congestion must not change sharply on or near October 14, 2003. The exact timing of the strike corresponded to the expiration of a 60-day court injunction and is thus exogenous.
- Nevertheless, it is important to rule out any possibility of seasonal effects influencing results, particularly since the strike began the first day following a three-day weekend

FALSIFICATION

Neighboring counties not covered by MTA

Time placebo- one year after the strike

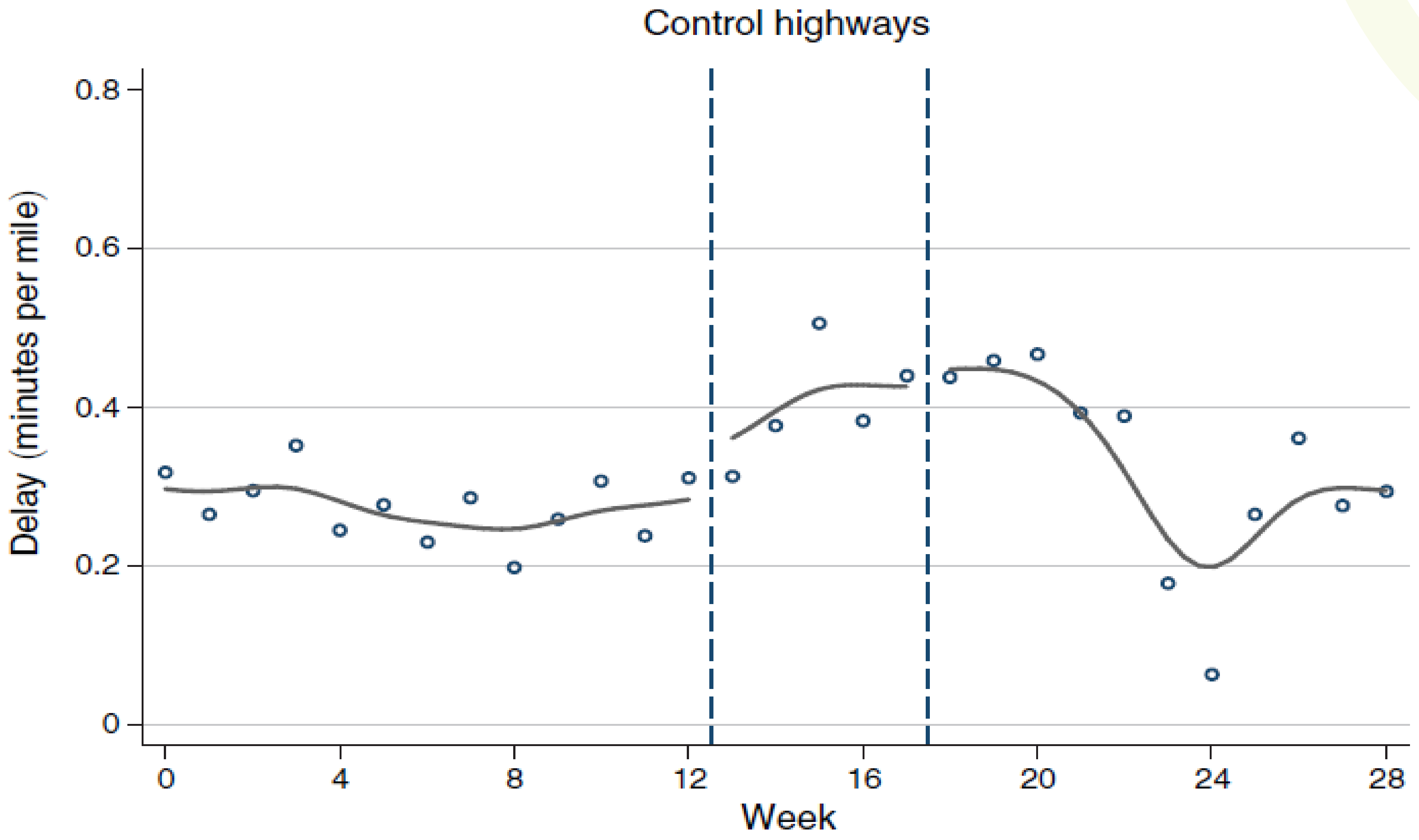


FIGURE 5. WEEKLY PEAK HOUR DELAY ON ORANGE/VENTURA COUNTY FREEWAYS, 7/14/2003 TO 1/30/2004

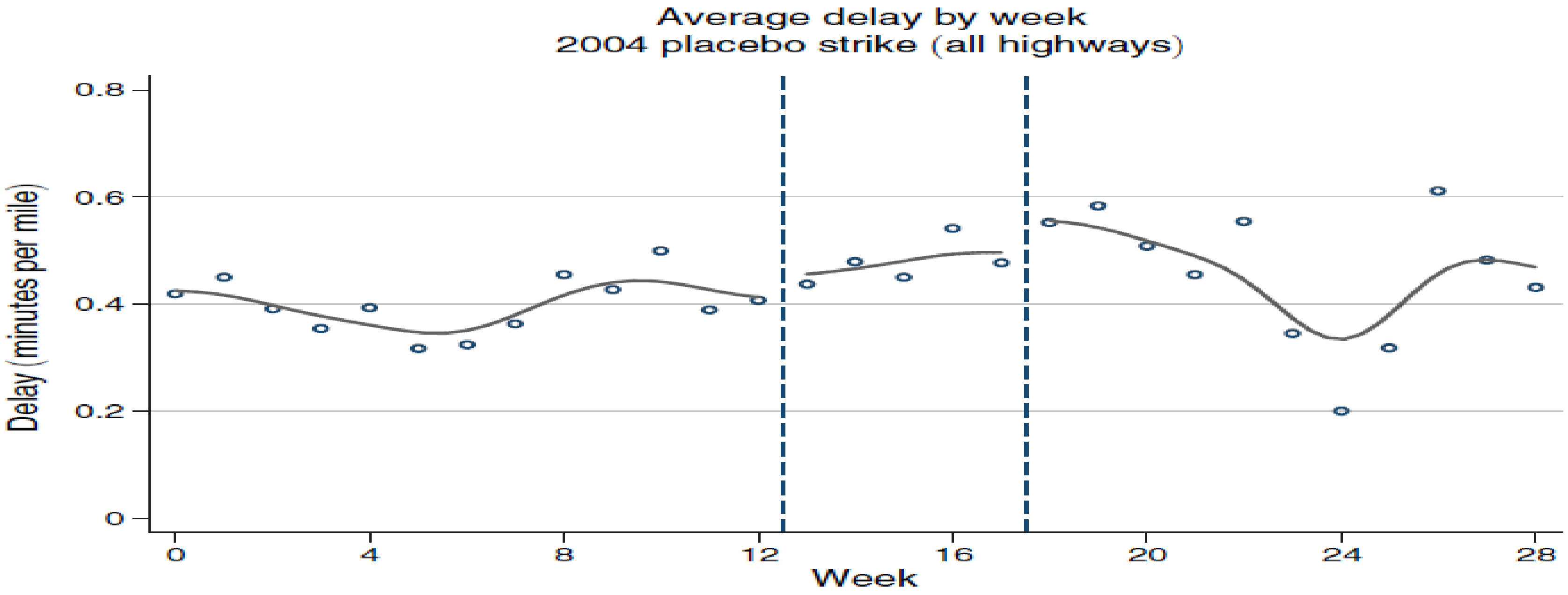


FIGURE 6. WEEKLY PEAK HOUR DELAY ON MAJOR LOS ANGELES FREEWAYS ONE YEAR LATER, 7/14/2004 TO 1/30/2005

GASOLINE CONTENT REGULATION *AU FFHAMMER AND KELLOGG (2011)*

$$(3) \qquad \ln(y_{it}) = \alpha_i \cdot \mathbf{Treat}_{ct} + \beta_i \cdot \mathbf{W}_{it} + f_i(Date_t) + \mu_i + \varepsilon_{it}.$$