Regression Discontinuity Aggregation, with an Application to the Union Effects on Inequality

Kirill Borusyak
UC Berkeley

Matan Kolerman-Shemer
The Hebrew University of Jerusalem*

December 2024

Abstract

We extend the regression discontinuity (RD) design to settings where each unit's treatment status is an average or aggregate across multiple discontinuity events. Such situations arise in many studies where the outcome is measured at a higher level of spatial or temporal aggregation (e.g., by state with district-level discontinuities) or when spillovers from discontinuity events are of interest. We propose two novel estimation procedures — one at the level at which the outcome is measured and the other in the sample of discontinuities — and show that both identify a local average causal effect under continuity assumptions similar to those of standard RD designs. We apply these ideas to study the effect of unionization on inequality in the United States. Using credible variation from close unionization elections at the establishment level, we show that a higher rate of newly unionized workers in a state-by-industry cell reduces wage inequality within the cell.

*Borusyak: k.borusyak@berkeley.edu; Kolerman-Shemer: matan.kolerman@mail.huji.ac.il. We thank Shmuel San for helpful suggestions and Jacob Lefler for excellent research assistance.

1 Introduction

Regression discontinuity (RD) designs are a popular tool for causal inference. In canonical sharp RD designs, the treatment is determined by whether a "running variable" exceeds a fixed cutoff. For instance, one may study whether unionization affects establishment outcomes, leveraging the cutoff of 50% votes in a union election (DiNardo and Lee, 2004; Frandsen, 2021). A local causal parameter is typically estimated via local linear regression under the assumption that potential outcomes are continuous around the cutoff. For standard RD methods to apply, each unit has to be exposed to only one discontinuity.

In this paper, we extend the RD design to settings in which a unit is exposed to multiple discontinuity events because the outcome is defined at a higher level of aggregation (e.g., in space or time) than the elections.[1] The endogenous explanatory variable ("treatment") in such settings is typically a simple or weighted average or sum of the RD shocks across multiple elections. We propose two new estimators in these "RD aggregation" (RDA) settings. The "upper-level" estimator is based on an instrumental variable (IV) regression at the level at which the outcomes are measured, with a weighted average or sum of the RD shocks from narrow elections as the instrument, while controlling for similar weighted sums of the controls used in local linear estimation (e.g., of the running variables). The "lower-level," or "stacking," estimator is a fuzzy RD estimator in the sample of narrow elections, with the unusual feature that the same values of the outcome and endogenous variable are repeated for the elections affecting the same unit. We show that, under assumptions similar to those of standard RD, both estimators identify the same local average treatment effect. Moreover, Monte Carlo simulations confirm that these estimators inherit attractive bias-reduction properties of conventional RD methods. The setting we consider nests many empirical literatures, and our ideas yield practical insights for them. We then apply the proposed methods to estimate the effects of unionization on earnings inequality at the level of state-by-industry cells, with typical cells exposed to multiple establishment-level unionization events.

The primary use cases for the RDA approach are empirical settings in which RD events are aggregated across space, time, or in other ways to construct a treatment variable at the level at which the outcome is measured. Such settings are very common. First, a large group of papers considers legislatures that consist of politicians elected in single-seat constituencies. The researcher is interested in the impact of the seat share of politicians who possess a specific characteristic: e.g., gender (Clots-Figueras, 2011, 2012; Bhalotra and Clots-Figueras, 2014; Ben-Porath and Kolerman-Shemer, 2024), religion (Bhalotra et al., 2014, 2021), or affiliation with a specific party (Folke, 2014; Nellis et al., 2016; Nellis and Siddiqui, 2018; Ben-Porath and Kolerman-Shemer, 2024). For instance, a state-level outcome is regressed on the overall seat share of women; the seat share of women who won close races is used as an IV. Second, in many cases a unit can be exposed to multiple RD events over time (Cellini et al. (2010); Dell and Querubin (2018); Biasi et al. (2024)). In the RDA approach, the researcher can specify the treatment as the total fraction of the period during which the unit is treated. Other studies aggregate discontinuities across investors in a region, across university courses taken by a student, across politicians that a firm is linked to, across destinations that a city is connected to by a direct flight, and across unionization elections to which we return below.[2]

[1] We use the term "elections" broadly to refer to events that generate RD variation, as the majority of RD papers we cite are indeed about elections of some sort.

[2] Specifically, Ater et al. (2021) construct a regional housing supply shock by aggregating across real estate investors. They leverage a discontinuity in the capital gains tax law that pushes some investors to sell their property. Tan (2023) considers the effects of the number of high letter grades of a university student, as aggregating across individual courses. Akey (2015) studies firms that donate to multiple politicians and measures the effects of a firm's political connectedness, as aggregating election outcomes of those politicians. Campante and Yanagizawa-Drott (2018) aggregate discontinuities in flight availability between city pairs with respect to distance to construct a shock to a city's overall connectivity. Kolerman-Shemer (2023) and

Our results also extend naturally to estimating spillover effects from discontinuity events, such that each unit is exposed to multiple elections of their neighbors (in a spatial or network sense). Isen (2014), for instance, leverages close referenda on tax changes in Ohio counties to study the effects on neighboring counties. Similarly, Dechezleprêtre et al. (2023) study the effects of tax breaks, which have discontinuous eligibility rules, on technologically similar firms.

While the two estimators we propose are natural extensions of conventional RD to RDA settings and they possess attractive statistical properties, they differ from most *ad hoc* estimation approaches used in the existing empirical literature. Specifically, papers estimating upper-level specifications do not include all local linear controls aggregated in the way suggested by our upper-level solution. As a result, they generally forfeit the reduced bias advantages of local linear estimation that made it standard in conventional RD designs. Conversely, studies using lower-level specifications include local linear controls but, in many cases, use reduced-form estimation which does not define a coherent causal model for the upper-level outcome. Some of these papers also impose restrictions limiting the sample that is not necessary with the RDA approach. Finally, some papers resort to difference-in-differences and other non-RD methods in settings where RD would be a standard tool for credible identification, if not for the aggregated outcomes. The RDA approach makes RD identification feasible in those settings.

We apply the RDA approach to study the union effects on earnings inequality. This question is one of the oldest in labor economics, dating back to at least Lewis (1963). It is also of central policy importance in modern days: Barack Obama, for instance, noted that *"as union membership has fallen, inequality has risen"* (Obama, 2016)— a correlation observed in the data by Card (2001). Yet, establishing a causal link between unionization and inequality has proved difficult. Recently, Fortin et al. (2022) point out that *"empirical evidence on the distributional impact of [union] threat effects is limited by the challenge of finding exogenous sources of variation in the rate of unionization."* Similarly, Farber et al. (2021) for the most part use a selection-on-observables designs. We make progress by leveraging close union elections in an RDA design.

We find that a higher rate of unionization leads to a decline in various measures of inequality at the level of state-by-industry cells: the Gini index, the 90/10 ratio, the share of labor income by top 10% earners, and the variance in log wages. These impacts result primarily from the decline of wages at the top of the wage distribution rather than an increase at the bottom. While this result calls for further investigation, we find evidence for one mechanism that can help explain it: pension coverage increases in cells with higher unionization rates. In a back-of-the-envelope calculation, we find that around 35% of the growth in the key within-cell inequality measures between 1970 and 2010 can be explained by the declining rates of unionization relative to the 1960s. Our estimate of the union contribution is moderately larger than the most comparable estimates from Farber et al. (2021).

We believe our empirical strategy for estimating the impacts of unionization on aggregated outcomes can be helpful beyond the effects on inequality. Some papers evaluated such effects by comparing neighboring counties across the border of right-to-work laws, which were mostly passed in the 1940s and 1950s and weakened unions: Holmes (1998), Feigenbaum et al. (2019), and Makridis (2019) studied the impacts on productivity, manufacturing activity, and voting, respectively. Other papers leveraged specific case studies that generate exogenous variation in union strength when looking at outcomes such as wages (Biasi, 2021) and the gender pay gap (Biasi and Sarsons, 2022). In contrast, the RDA method allows researchers to leverage credible RD variation from multiple decades and the whole common, which is standard when studying the impacts on establishment-level outcomes. Complementary work by Kolerman-Shemer (2023) has used the

the application in this paper aggregate establishment-level unionization data from close elections to estimate the effects of unionization on political outcomes and inequality, respectively.

RDA method to show that unionization makes local politics more Democratic in the US.

Methodologically, our paper bridges the literatures on shift-share instruments and RD designs. Our theoretical analysis rests on the observation that both the treatment and the instrument in the RDA setting are shift-share aggregates (i.e., weighted sums) of RD shocks. We then leverage the results from Borusyak et al. (2022) to show how upper-level IV regressions with the aggregated instrument and controls are numerically equivalent to particular fuzzy RD regressions at the lower level. This equivalence is useful in two ways: on one hand, it brings upper-level regressions closer to standard fuzzy RD and motivates our choice of upper-level controls as aggregations of the standard controls from local linear estimation. On the other hand, it motivates our lower-level estimator by showing that an appropriate lower-level regression can identify a parameter at the upper level, at which the causal effect is naturally defined.

The setting we study is distinct from "multi-score" RD settings which also feature multiple running variables per unit (see Cattaneo and Titiunik (2022) for a review). For instance, in spatial RD designs (e.g., Black, 1999; Dell, 2010), the latitude and longitude are the two running variables, and a geographic boundary separates the treated and control areas. Another example is when a student is treated if several tests scores (e.g., for math and English) simultaneously exceed some cutoffs. Such designs feature a single binary treatment (or instrument, in the fuzzy case), and a simple transformation of the running variables reduces the problem to a standard RD, in which the signed distance to the boundary is the scalar running variable. In contrast, in our setting the instrument aggregates RD shocks. While still a function of the underlying running variables, it does not feature a single boundary, making existing methods inapplicable.

A recent literature has examined certain RD settings with interference. Torrione et al. (2024) extend the multi-score approach, based on the minimum Euclidian distance from the set of neighbors' running variables to the boundary at which the spillover treatment, e.g. the number or share of treated neighbors, switches between two prespecified values. This allows them to identify more detailed spillover effects, e.g. of having all treated neighbors vs. all untreated neighbors conditionally on the unit itself being treated. We propose a very different estimation approach which identifies a causal spillover parameter in a simpler way, does not depend on the discreteness of the treatment, and applies to a variety of settings beyond just spillovers. Other papers have examined the case of interference occurring among units with similar values of the running variable, as may be natural in spatial RD designs (Keele and Titiunik (2018), Auerbach et al. (2024)). In our setting, even if the structure of spillovers is based on geography, each unit has its own running variable, leading to different issues and solutions.

While our RDA design considers treatments and instruments that are shift-share aggregates of RD shocks, a recent literature has considered more complex instruments constructed from RD variation, via the idea of recentering. As Borusyak and Hull (2023) pointed out, if within some bandwidth around the cutoff the RD shocks are viewed as completely random (as in the local randomization approach to RD; see Cattaneo et al. (2015)), one can compute the expectation of the instrument over the distribution of RD shocks and control for it to isolate the exogenous variation. Kott (2022) and Narita and Yata (2023) compute the expected instrument by simulating counterfactual realizations of RD shocks, while Abdulkadiroğlu et al. (2022) and Angrist et al. (2024) compute it analytically. Our paper focuses on the settings where the instrument is linear in the RD shocks — which still covers a wide variety of empirical applications. Thanks to this focus, we are able to leverage the local linear estimation approach to RD and develop estimators with better statistical properties than those based on local randomization.

The rest of this paper is organized as follows: in Section 2 we present our theoretical framework and results in the context of cross-sectional RD aggregation, as well as Monte Carlo simulations for the proposed

estimators. In Section 3, we discuss how our proposals differ from the current empirical practice. In Section 4, we extend the analysis to settings with spillovers and temporal aggregation. In Section 5, we apply the RDA approach to estimate the impact of unions on US inequality. Section 6 concludes.

2 Theoretical Framework

In Section 2.1, we introduce the setting and compare it with conventional RD designs. Sections 2.2 and 2.3 propose two estimators that we put forward and provide intuition why they may provide valid causal estimates. Section 2.4 formally shows how, under conditions similar to those of standard fuzzy RD, these estimators identify a local average causal effect, and Section 2.5 establishes their attractive bias-reduction properties in a Monte Carlo simulation.

2.1 Setting

We measure outcome Y_i for a set of units $i = 1, \ldots, N$, which we refer to as the "upper level." Each unit consists of a set $\mathcal{J}_i$ of "lower-level" subunits j (with $|\mathcal{J}_i| = J_i > 0$). A subunit is characterized by a running variable r_j and treatment $z_j = \mathbf{1}[r_j \geq 0]$ that happens whenever the running variable exceeds the cutoff normalized to zero.[3] We are interested in estimating the causal model

$$Y_i = \beta X_i + \varepsilon_i, \tag{1}$$

where ε_i is the untreated potential outcome and β is the causal effect (assumed homogeneous for now) of some upper-level treatment X_i. Our benchmark case is when X_i simply aggregates z_j across all relevant subunits,

$$X_i = \sum_{j \in \mathcal{J}_i} s_j z_j, \tag{2}$$

but other treatments can be considered, too (see, e.g., Section 4.1 for an example). Here $s_j > 0$ are known importance weights which may or may not be the same for all $j \in \mathcal{J}_i$ and may or may not add up to one.[4] In the election example, X_i may be the fraction of seats in the state i legislature represented by candidates from the focal party. Here z_j indicates the focal party's victory in district j and s_j measures the seat share of the district j in the legislature. In the common case of one seat per district, s_j is the inverse of the total number of seats.

The conventional sharp RD setting corresponds to the case where the outcomes are measured at the same level as the running variables, i.e. i and j are the same objects: $\mathcal{J}_i = \{i\}$, $s_j = 1$, and $X_i = z_i$. In that scenario, the standard local linear estimation approach restricts the sample to observations near the cutoff ($|r_j| \leq h$ for bandwidth h) and estimates by ordinary least squares (OLS) the specification

$$Y_i = \beta z_j + \gamma_0 + \gamma_1 r_j + \gamma_2 r_j^+ + \text{error}_j, \qquad j \equiv i, \tag{3}$$

where the controls $q_j = (1, r_j, r_j^+)$ include the intercept, the running variable, and its interaction with being on the right of the cutoff (i.e., $r_j^+ = r_j \cdot \mathbf{1}[r_j \geq 0]$). Sometimes this regression is weighted by "kernel" weights

[3] We use capital and small letters to denote unit- and subunit-level variables, respectively.

[4] We assume that the sets are non-overlapping to simplify notation. The setting naturally extends to the case where different units can be exposed to the same subunits, as when estimating the spillovers of RD events; see Section 4.1.

that are larger when r_j is closer to the cutoff; importance weights are also allowed.

But how can local variation in r_j be leveraged to obtain a consistent estimate of β in (1) when each unit can be exposed to multiple subunits, potentially with more than one near the cutoff? We next propose two strategies, one at the upper level and the other at the lower level, focusing on the intuitions before proceeding to formal and simulation-based analysis of consistency and bias.

2.2 An Upper-Level Solution

Our first proposal, at the level of upper-level of observations i, involves instrumenting X_i with its component arising from the subunits whose running variables are near the cutoff:

$$Z_i = \sum_{j \in \mathcal{C}_i} s_j z_j, \tag{4}$$

where $\mathcal{C}_i = \{j \in \mathcal{J}_i \colon \mathbf{1}[|r_j| \leq h]\}$ is the set of i's subunits near the cutoff. In the election example, Z_i is the fraction of state legislature's seats represented by a candidate from the focal party who won in a close election, out of all seats.

We further propose to aggregate the vector of standard RD controls q_j to the i level in the same way as the instrument is constructed,

$$Q_i = \sum_{j \in \mathcal{C}_i} s_j q_j = \big(\sum_{j \in \mathcal{C}_i} s_j, \sum_{j \in \mathcal{C}_i} s_j r_j, \sum_{j \in \mathcal{C}_i} s_j r_j^+\big)'. \tag{5}$$

We label these three aggregated variables the *RDA controls*. The first of them is the total weight of subunits near the cutoff. For instance, in the election example, it is the fraction of legislature seats determined in a close election. The second is the average running variable in narrow elections, rescaled by their total weight. And the last one is the average running variable in narrow wins, rescaled by the total weight of those narrow wins.

As a result, the upper-level RDA estimator corresponds to the instrumental variable (IV) specification

$$Y_i = \beta X_i + \gamma_0 \sum_{j \in \mathcal{C}_i} s_j + \gamma_1 \sum_{j \in \mathcal{C}_i} s_j r_j + \gamma_2 \sum_{j \in \mathcal{C}_i} s_j r_j^+ + \tilde{\gamma}' \tilde{W}_i + \text{error}_i, \tag{6}$$

instrumented by Z_i, where $\tilde{W}_i$ includes the i-level intercept and possibly additional variables that are continuous around the cutoff. Its reduced-form parallels (3):

$$Y_i = \rho \sum_{j \in \mathcal{C}_i} s_j z_j + \gamma_0 \sum_{j \in \mathcal{C}_i} s_j + \gamma_1 \sum_{j \in \mathcal{C}_i} s_j r_j + \gamma_2 \sum_{j \in \mathcal{C}_i} s_j r_j^+ + \tilde{\gamma}' \tilde{W}_i + \text{error}_i. \tag{7}$$

While aggregating both the instrument z_j and the covariates q_j in the same way is intuitive, why do we expect RDA to identify a causal parameter? We provide intuition by showing that the IV estimator from (6) is numerically equivalent to a standard fuzzy RD estimator from a transformed regression, and can therefore be expected to inherit some its properties.

Indeed, Z_i can be viewed as a shift-share instrument constructed from shifts z_j and exposure shares s_j.[5] We can therefore leverage the numerical equivalence result of Borusyak et al. (2022, Prop. 1) to represent $\hat{\beta}$

[5] More precisely, the shares are s_j for $j \in \mathcal{C}_i$ and zero for $j \notin \mathcal{C}_i$. While shift-share instruments usually involves shifts that are common to all observations, this is not a requirement (Borusyak et al., 2024, Appendix A.6).

as an IV estimator at the subunit level:

Proposition 1. *For any unit-level variable V_i let $V_i^{\perp}$ denote the residual from a projection of V_i on $W_i = \left(Q_i, \tilde{W}_i\right)'$. Let $\boldsymbol{i}(j)$ denote the unit to which subunit j belongs. Then $\hat{\beta}$ from (6) equals the IV estimator of β from a subunit-level specification*

$$Y_{\boldsymbol{i}(j)}^{\perp} = \beta X_{\boldsymbol{i}(j)}^{\perp} + \lambda' q_j + error_j, \tag{8}$$

on the sample of $j \in \mathcal{C} \equiv \cup_i \mathcal{C}_i$, where $X_{\boldsymbol{i}(j)}^{\perp}$ is instrumented by z_j and the regression is weighted by s_j.

Proof. Follows from Proposition 1 in Borusyak et al. (2022) by observing that the shock-level weight defined in that paper, $\sum_i \sum_{j \in \mathcal{C}} s_j \mathbf{1}[j \in \mathcal{C}_i]$, simplifies to s_j, and that shock-level averages simplify to $\bar{V}_j = \sum_i s_j \mathbf{1}[j \in \mathcal{C}_i] V_i / \sum_i s_j \mathbf{1}[j \in \mathcal{C}_i] = V_{\mathbf{i}(j)}$. □

The specification (8) is artificial but it is useful because it maps directly to the standard fuzzy RD design: the instrument is the subunit-level RD shock z_j and the covariates are $q_j = (1, r_j, r_j^+)$.[6] One can use the equivalent j-level representation of the upper-level specification to choose bandwidth and to perform statistical inference on the estimator, as in Calonico et al. (2014) .

2.3 RD Stacking: A Lower-Level Solution

Our second proposal is to estimate a fuzzy RD specification on the sample that stacks all subunits near the cutoff:

$$Y_{\mathbf{i}(j)} = \beta X_{\mathbf{i}(j)} + \tilde{\gamma}' \tilde{W}_{\mathbf{i}(j)} + \lambda' q_j + \text{error}_j, \qquad j \in \cup_i \mathcal{C}_i \tag{9}$$

with $X_{\mathbf{i}(j)}$ instrumented by the subunit-level treatment z_j and using s_j as importance weights. Kernel weights may be additionally used, too.[7] Since (9) is a conventional fuzzy RD specification, it is clear it uses local variation around the cutoff.

The specification (9) is peculiar: it provides multiple expressions for Y_i of the same unit. One may also find it odd that the instrument z_j varies among multiple observations — subunits of the same unit — that mechanically have the same value of the treatment $X_{\mathbf{i}(j)}$.

Despite this peculiarity, Proposition 1 demonstrates why specification (9) may be helpful. It only differs from specification (9) by having the outcome and treatment initially residualized on W_i. These controls, including Q_i, are continuous around the cutoff and therefore may be expected not to affect identification. Moreover, since local linear controls q_j enter both specifications, one may hope that both approaches inherit the bias properties of standard RD. Below we show that such hope is justified.

The stacking approach has an additional advantage over the upper-level solution: since it is just a fuzzy RD specification with raw outcome and treatment, it naturally lends itself to standard RD plots for the reduced-form and first-stage of (9).

2.4 Identification of a Local Causal Effect

To describe the local parameter identified by RDA, we introduce the following causal model with heterogeneous causal effects. For $\boldsymbol{z}_i = (z_j)_{j \in \mathcal{J}_i}$, let $X_i(\boldsymbol{z}_i)$ denote the potential treatment of unit i depending on

[6] While in many fuzzy RD designs the treatment is binary, this is not necessary (e.g., Almond et al., 2010, Section VIII).

[7] Kernel weights should not be used when constructing Z_i for the upper-level specification. Correspondingly, they were not introduced in Section 2.2.

the treatments of its subunits. When equation (2) holds, it defines $X_i(\boldsymbol{z}_i)$, but we allow for general X_i. We further make an exclusion restriction that $\boldsymbol{z}_i$ only affects Y_i via X_i, with $Y_i(x)$ denoting potential outcomes.[8] We further impose:

Assumption 1 (Sampling of units)**.** *The data are a random sample of* $J_i, \boldsymbol{r}_i, \boldsymbol{z}_i, \boldsymbol{s}_i, \boldsymbol{q}_i, X_i, X_i(\cdot), \tilde{W}_i, Y_i, Y_i(\cdot)$ *where bold symbols indicate* J_i*-dimensional vectors containing similarly named subunit-level variables, and consistency conditions hold:* $z_j = \mathbf{1}[r_j \geq 0]$*,* $X_i = X_i(\boldsymbol{z}_i)$*, and* $Y_i = Y_i(X_i)$*.*

Assumption 2 (Irrelevance of labels)**.** *For any* J *in the support of* J_i*, the distribution of* $(\pi(\boldsymbol{r}_i), \pi(\boldsymbol{z}_i), \pi(\boldsymbol{s}_i), \pi(\boldsymbol{q}_i), X_i, X_i(\pi(\cdot)), \tilde{W}_i, Y_i, Y_i(\cdot)) \mid J_i = \bar{J}$ *is the same for all permutations* π *of* $\bar{J}$ *elements.*

Under these assumptions, the distribution of unit-level data induces a distribution of the relevant variables over subunits, such as the joint distribution of z_j and $\boldsymbol{z}_{\mathbf{i}(j)-j}$ where $\left(z_j, \boldsymbol{z}_{\mathbf{i}(j)-j}\right) \equiv \boldsymbol{z}_{\mathbf{i}(j)}$. We then impose a generalization of the RD continuity assumption:

Assumption 3 (Continuity)**.** *(a) Expected reweighted potential treatment and outcome,* $\mathbb{E}\left[s_j X_{\boldsymbol{i}(j)}(z, \boldsymbol{z}_{-j}) \mid r_j = r\right]$ *and* $\mathbb{E}\left[s_j Y_{\boldsymbol{i}(j)}(X_{\boldsymbol{i}(j)}(z, \boldsymbol{z}_{\boldsymbol{i}(j)-j})) \mid r_j = r\right]$ *are continuous in* r *at* $r = 0$ *for* $z = 0, 1$*.*

(b) Expected importance weights $\mathbb{E}\left[s_j \mid r_j = r\right]$ *and the cumulative distribution function of* $\boldsymbol{z}_{\boldsymbol{i}(j)-j} \mid r_j = r$ *are continuous in* r *at* $r = 0$*.*

While this assumption is high-level, it will be satisfied when the distributions of raw potential outcomes $Y_{\mathbf{i}(j)}(x)$, potential treatments $X_{\mathbf{i}(j)}(z, \boldsymbol{z}_{\mathbf{i}(j)-j})$, importance weights, and leave-out treatments $\boldsymbol{z}_{\mathbf{i}(j)-j}$ are continuous around the cutoff for r_j. The last assumption is satisfied when the vector of running variables $\boldsymbol{r}_i$ has joint density, in particular ruling out the case where multiple running variables are perfectly correlated.[9]

Assumption 4 (Density)**.** *The density of* r_j *is positive and continuous at 0.*

Assumption 5 (Monotonicity and first stage)**.** $X_{\boldsymbol{i}(j)}(1, \boldsymbol{z}_{\boldsymbol{i}(j),-j}) \geq X_{\boldsymbol{i}(j)}\left(0, \boldsymbol{z}_{\boldsymbol{i}(j),-j}\right)$ *a.s. and* $\mathbb{E}\left[s_j \left(X_{\boldsymbol{i}(j)}(1, \boldsymbol{z}_{\boldsymbol{i}(j),-j}) - X_{\boldsymbol{i}(j)}(0, \boldsymbol{z}_{\boldsymbol{i}(j),-j})\right) \mid r_j = 0\right] > 0$*.*

Assumption 5 is trivially satisfied when the treatment is defined by equation (2). For simplicity, we abstract away from additional controls, setting $\tilde{W}_i = 1$; Calonico et al. (2019) show that adding predetermined covariates in a linearly separable way does not affect the RD estimand. The following proposition shows that under these assumptions both of our proposals identify the same local average treatment effect which puts more weight on units with more numerous and more important subunits near the cutoff, as well as subunits with a stronger first-stage impact on the treatment.

Proposition 2. *Suppose Assumptions 1–5 hold. Then, as* $h \to 0$*, the estimand* $\beta_h^u \equiv \mathrm{Cov}\left[Y_i^{\perp}, Z_i\right] / \mathrm{Cov}\left[X_i^{\perp}, Z_i\right]$ *of (6) and the similarly defined estimand* β_h^{ℓ} *of (9) converge to the same convexly-weighted average of potential outcome slopes:*

$$\beta_0 \equiv \frac{\mathbb{E}\left[s_j \cdot \left(Y_{\boldsymbol{i}(j)}\left(X_{\boldsymbol{i}(j)}(1, \boldsymbol{z}_{\boldsymbol{i}(j)-j})\right) - Y_{\boldsymbol{i}(j)}\left(X_{\boldsymbol{i}(j)}\left(0, \boldsymbol{z}_{\boldsymbol{i}(j)-j}\right)\right)\right) \mid r_j = 0\right]}{\mathbb{E}\left[s_j \cdot \left(X_{\boldsymbol{i}(j)}(1, \boldsymbol{z}_{\boldsymbol{i}(j)-j}) - X_{\boldsymbol{i}(j)}\left(0, \boldsymbol{z}_{\boldsymbol{i}(j)-j}\right)\right) \mid r_j = 0\right]}. \tag{10}$$

Proof. See Appendix A. □

[8] Here we assume for clarity that both the first and second stages of (6) are causal. Both assumptions can be relaxed as in Small et al. (2017) and Appendix A.1 of Borusyak et al. (2022).

[9] However, when multiple running variables are perfectly correlated, they can be combined into one without loss.

We note that the proof makes it clear that, among the covariates, only $\sum_{j\in\mathcal{C}_i} s_j$ in the upper-level specification and the intercept in the lower-level specification are necessary for identification. A parallel to the local randomization approach to RD designs provides intuition for this conclusion: if the running variables are viewed as drawn from a uniform distribution, controlling for $\sum_{j\in\mathcal{C}_i} s_j$ is equivalent to controlling for the expected value of Z_i, as in the Borusyak and Hull (2023) recentering procedure. This control isolates a unit's "luck" in getting a higher-than-expected vs lower-than-expected value of the instrument. While our identification result (or the local randomization approach) do not provide a justification for including the aggregated local linear controls, the next subsection shows that they are important for reducing bias, as in standard RD designs based on the continuity assumption.

We also note that the effective weights on causal effects in (10) diverge from the importance weights the regressions are estimated with. Consider the case where X_i is defined as in equation (2), such that the first-stage effect of z_j equals s_j, and the causal effects are heterogeneous across units but linear in x, $Y_i(x) = Y_i(0) + \beta_i x$. Then the total weight placed on β_i in (10) is $\sum_{j\in\mathcal{C}_i} s_j^2$. When $s_j = 1/J_i$, as when the treatment is the *share* of legislature seats won by certain candidates (e.g., female or Democratic), the effective weight is smaller in larger legislatures, as the law of large numbers eliminates much of the variation in the instrument. If, however, $s_j = 1$, as when the treatment is the *number* of such legislature seats, the effective weight is larger in larger legislatures, as they will tend to have more individual seats with narrow elections. Importance weights in both upper- and lower-level regressions can be adjusted if other averages of causal effects are of interest.

2.5 Bias and Efficiency: Monte Carlo Simulations

The key advantage of the two estimators we propose is that they inherit the attractive bias and variance properties of standard RD estimators. In particular, when the bandwidth shrinks to zero, their bias shrinks at a quadratic rate or faster. We leave theoretical results establishing this property to future drafts; for now, we illustrate the performance of alternative estimators via a Monte Carlo simulation.

We report the performance of the upper-level IV estimator with specification (6) and the lower-level IV estimator with specification (9).[10] We contrast them with a benchmark upper-level IV estimator which includes the control necessary for identification, $\sum_{j\in\mathcal{C}_i} s_j$, but not the other controls aimed at reducing bias. Section 3 we show that the latter estimator is commonly used in practice, and it also aligns with the recentering approach based on the local randomization view of RD.

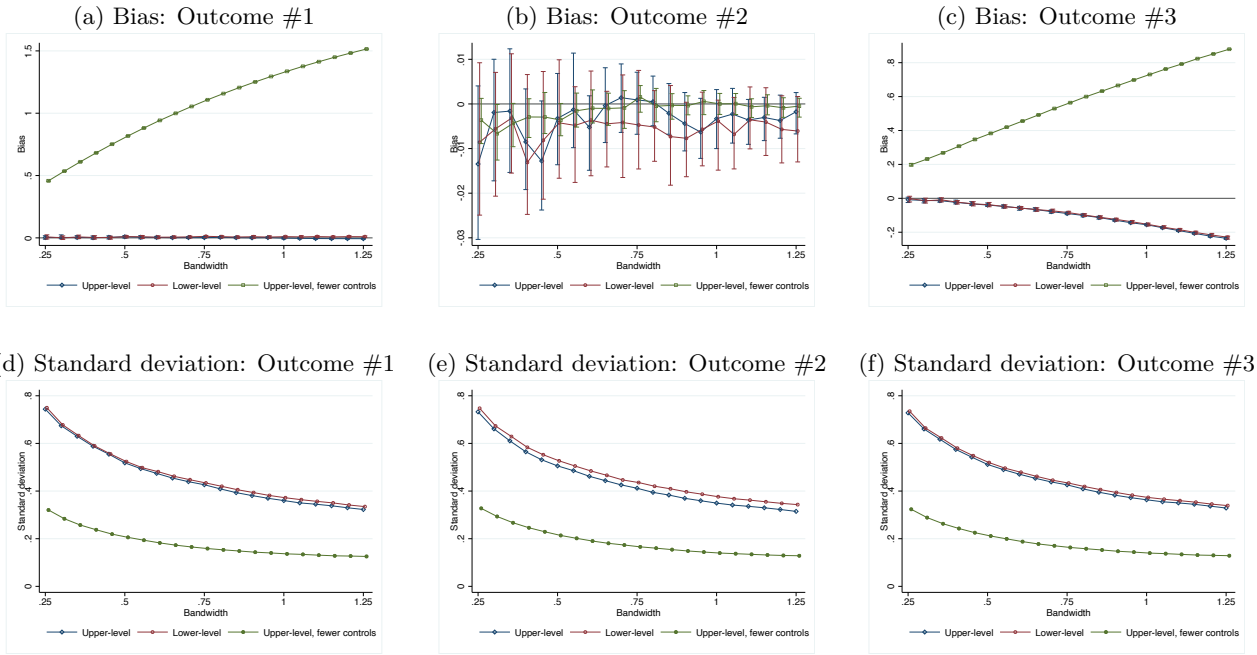
In each simulation, we draw 1,000 units (cells), each with 5 subunits (elections) with equal importance weights $s_j = 1/5$. For each subunit, we generate the running variable as $r_j = \rho\zeta_{\mathbf{i}(j)} + \sqrt{1-\rho^2}\cdot\xi_j$, where ζ_i and ξ_j are i.i.d. standard normal variables and $\rho = 0.5$ captures the correlation of running variables among subunits of the same unit. The treatment variable is defined as the share of treated subunits, $X_i = \sum_{j\in\mathcal{J}_i} s_j z_j$ for $z_j = \mathbf{1}[r_j > 0]$. To focus on the bias and variance, we assume the true causal effect is zero for all observations.

The top row of Figure 1 reports the median bias of the three estimators for a range of bandwidth values, along with the bootstrap 95% confidence interval for this median.[11] The bias depends crucially on how potential outcomes vary with the running variables of the subunits. We therefore consider several data-generating processes for the outcome; focusing here on the median bias, we do not include an independent

[10] We do not include any additional controls besides the intercept, setting $\tilde{W}_i = 1$.

[11] The range $h \in [0.25, 1.25]$ corresponds to approximately $[20\%, 80\%]$ of subunits included in the close-elections sample. The confidence interval is hard to see on some panels because the number of simulations is large enough for the median to be very precisely estimated.

Figure 1: Monte Carlo Simulations



Notes: The top row reports, for each bandwidth value, the median bias across 2,500 simulated datasets, along with the 95% bootstrap confidence interval for it based on 300 bootstrap draws. The bottom row reports standard deviations of estimates across simulations. See main text for the data-generating processes and definitions of the estimators.

error. Panel (a) considers a unit-level outcome that is linear in the running variables of all subunits, $Y_i = \sum_{j \in \mathcal{J}_i} s_j r_j$. In standard RD, that would eliminate bias completely, as long as the local linear controls are included. In the RDA setup, this is not the case, as the same unit can include both close and non-close elections. Yet, the qualitative behavior is the same: the bias of our proposed upper- and lower-level estimators is negligible compared to the benchmark upper-level estimator with fewer controls (and the 95% bootstrap confidence interval includes zero for most bandwidths).[12] Panel (b) considers an outcome which depends on the running variables nonlinearly but with the same curvature on either side of the cutoff, $Y_i = \sum_{j \in \mathcal{J}_i} s_j r_j^2$. Here all three estimators achieve negligible bias; note the scale of the vertical axis as well as the 95% confidence interval for the bias estimate based on our 2,500 simulations. Finally, panel (c) considers the case where curvature changes around the cutoff, $Y_i = \sum_{j \in \mathcal{J}_i} s_j r_j^2 \mathbf{1}[r_j > 0]$. This is the case studied in the standard RD setup by, e.g., Imbens and Kalyanaraman (2012) who show the importance of local linear controls in achieving quadratic decay of the bias. Inheriting this property, the bias of our proposed estimators is not only smaller than that for the benchmark one for all bandwidths; its slope also approaches zero for smaller h, which does not happen for the benchmark estimator.

In unreported simulations, we check robustness of these findings. First, we generate heterogeneous importance weights across subunits of the same unit (adding up to one). Second, while all outcomes in our main simulation depend symmetrically on all running variables, we also consider the outcome which equals the running variable of one arbitrarily picked subunit. Third, we increase the number of units from 1,000 to 2,000 or the number of subunits per unit from 5 to 10. The bias patterns are unchanged.

The results so far have shown that, at least for the data-generating processes we considered, the bias

[12] Another way to see this is that the reduced-form of our lower-level IV regression is exactly unbiased, as $\mathbb{E}[Y_i \mid r_j]$ is linear in this simulation.

of the upper- and lower-level regressions is essentially identical. There are, however, some differences in efficiency. To assess them, we repeat the simulation for the same outcomes adding *i.i.d.* standard normal noise to capture the idiosyncratic component of the error. The second row of Figure 1 reports standard deviations of the estimators across simulations. It shows that the upper-level regression is more efficient than the lower-level one, albeit the difference is small. The intuition is that aggregated controls better capture the correlation between potential outcomes and the running variables. This is not an artifact of our outcome definitions; rather, it follows from the irrelevance of labels (Assumption 2): to the extent the running variable of one subunit predicts the unit outcome, the appropriate average of them is at least as predictive.

To sum up, the Monte Carlo simulations confirm that including the aggregated local linear controls or estimating the effects at the lower-level in a standard fuzzy RD both succeed at substantially reducing bias. The upper-level estimator is also slightly more efficient, as it is better able to reduce residual variance.

3 Common Practices

Many published articles feature designs suitable for RDA. We identified over 50 such papers listed in Appendix Table A1. In this section, we describe the empirical strategies employed by the authors of those articles. We detail the main differences between their approaches and ours, highlighting the potential gains from using the estimators from Section 2. Section 3.1 discusses studies that, in the absence of the RDA methodology, did not use RD variation at all. We then turn to papers that used some RD specifications at the aggregated or disaggregated levels in Sections 3.2 and 3.3, respectively. Throughout the discussion, we generically refer to the subunits as elections.

3.1 Difference-in-Differences Specifications in RDA Settings

The first group of studies, listed in Panel A of Appendix Table A1, uses difference-in-differences and similar designs in settings where RDA estimation is feasible and could provide more credible identification. While the parallel trends assumption underlying difference-in-differences analyses is often hard to justify *ex ante*, continuity assumptions underlying conventional RD and inherited by RDA are usually more convincing.

Besley and Case (2003, Tables 8–10) consider the effects of legislature composition — one of the prime applications of RDA. To estimate the effects on policy choices of the fractions of female and Democratic legislators in U.S. state houses, where most seats are elected in single-seat races, they employ the two-way fixed effects OLS specification with a time-varying treatment:

$$Y_{it} = \beta X_{it} + \tilde{\gamma}' \tilde{W}_{it} + FE_i + FE_t + \text{error}_{it}.$$

A credible instrument for X_{it} could be obtained by leveraging the structure of the treatment as aggregating discontinuity events.

Similar examples are found in difference-in-differences studies of the effects of political connections on industries and firms. Akcigit et al. (2023, Table 10) estimate how a larger fraction of politically connected firms in an industry — defined as firms whose employee has been elected in a local council — affects industry outcomes. They use OLS estimation, despite using conventional RD for firm-level outcomes. Cooper et al. (2010) similarly consider the association between the number of politicians supported by a firm who get elected and firm performance.

Finally, Saez et al. (2019) estimate the wage effects of a payroll tax cut that applied to workers below the

age of 25. Their difference-in-differences specification compares firms with different shares of young workers, under a parallel trends assumption. The RDA approach would instead leverage the age cutoff, recognizing that a firm's share of young workers is an aggregate of worker-level discontinuities. The RDA strategy would thus entail comparing firms with workers age 25 or just above to firms with workers just below 25, providing credible RD variation.

3.2 Upper-Level Outcome and Upper-Level Specification

We now consider 19 papers from Panel B of Appendix Table A1 where RD variation is used in some way and the regression analysis is conducted at the upper-level (i-level in the notation of Section 2). These specifications can generally be represented as

$$Y_i = \beta X_i + \tilde{\psi}'\tilde{Q}_i + \tilde{\gamma}'\tilde{W}_i + \text{error}, \tag{11}$$

where X_i is constructed from all subunits and is instrumented by $Z_i = \sum_{j\in\mathcal{C}_i} s_j z_j$ constructed from the subunits near the cutoff only. Furthermore, $\tilde{Q}_i$ represents controls constructed in some way from subunit-level variables and $\tilde{W}_i$ are other controls and fixed effects.

The key difference between our proposal and what is found in the literature is in the set of controls $\tilde{Q}_i$. We discuss the three RDA controls Q_i in equation (6) in turn. The majority of papers include the first control: the total weight of narrow elections.[13] One exception is Valentim and Dinas (2024) who instead control for the weight of narrow elections *below* the cutoff; this produces a biased coefficient.[14]

The second RDA control — the aggregated running variable in narrow elections — is not found in any of the papers we have reviewed. While some papers have no analog to it (e.g., Campante and Yanagizawa-Drott (2018); Nellis and Siddiqui (2018)), several variations have been employed. Azoulay et al. (2019) take an unweighted average of running variables near the cutoff, while the instrument is a sum of RD shocks weighted by some dollar amounts. The same is true for Folke (2014) and Meriläinen (2022) whose average additionally includes non-close elections. Clots-Figueras (2012) and Bhalotra et al. (2014, 2021) use a different strategy: they control for the running variables of each subunit (including with non-narrow elections) separately. This requires ordering the subunits in an arbitrary way (such that changing the order would change the IV estimate) and filling missing values as zeros for units that have fewer subunits than others. Neither of these strategies appears to have a clear justification.

Finally, we are not aware of any study that included any version of the running variable aggregated among narrow wins only. Absent the second and third RDA variables, the estimators used in the literature do not inherit the properties of standard RD and therefore may not enjoy the low bias advantages of local linear estimation.

We conclude by noting that, while most of the papers estimate (11) by IV, Akey (2015) considers its reduced-form instead. We see several disadvantages to this approach. It is a less coherent specification than IV since the explanatory variable depends on the bandwidth choice, and thus is not an economic variable with independent meaning. Practically speaking, if the fraction of narrow victories is correlated with the fraction of non-narrow victories (as may happen when the bandwidth is not infinitesimal), the reduced-form

[13] Instead of controlling for it, some papers (e.g., Folke (2014), Azoulay et al. (2019)), in the language of Borusyak and Hull (2023), recenter the instrument Z_i by subtracting 0.5 times the total weight of narrow elections.

[14] To see this bias, notice that $\beta\sum_{j\in\mathcal{C}_i} s_j z_j + \gamma_0 \sum_{j\in\mathcal{C}_i} s_j = (\beta+\gamma_0)\sum_{j\in\mathcal{C}_i} s_j z_j + \gamma_0 \sum_{j\in\mathcal{C}_i} s_j(1-z_j)$. The left-hand side here corresponds to (6), while the right-hand side to the specification of Valentim and Dinas (2024). The reduced-form coefficient therefore identifies $\beta+\gamma_0$ instead of γ_0. Akey (2015) (Table 7, column 2) employs a similar specification, except he interprets the second coefficient as a causal effect of losing.

specification is biased. And even if the correlation is small or absent, the reduced-form specification is likely less efficient, as the error term includes the effects of non-narrow victories.

3.3 Upper-Level Outcome and Lower-Level Specification

The final group of 8 papers (Panel C of Appendix Table A1) considers outcomes measured at the more aggregated level than the discontinuity events but estimates certain regressions at the lower level. To structure the discussion, recall that the stacking approach of Section 2.3 involves the sample of all narrow elections and a fuzzy RD specification with the reduced-form

$$Y_{\mathbf{i}(j)} = \rho z_j + \lambda' q_j + \text{error}, \tag{12}$$

with $X_{\mathbf{i}(j)}$ as the endogenous variable in the structural equation, and s_j as weights. The current practice deviates in terms of the sample restrictions and the use of OLS vs IV.

First, while some papers (e.g., de Benedictis-Kessner and Warshaw (2020); de Benedictis-Kessner et al. (2024)) keep the full stacked sample of narrow elections, others restrict the sample in different ways. Beach and Jones (2017) only keep one narrow election for each unit, which has the running variable closest to the cutoff (although this does not affect too many units in their setting). Dell and Querubin (2018, Section III) only keep the narrow election that happened chronologically first. These approaches limit the amount of identifying variation, which our paper shows not to be necessary.

Second, while some papers (Carozzi et al. (2022) and Dell and Querubin (2018)) clearly introduce the upper-level treatment and use the IV specification (9), the majority of papers estimate the reduced-form (12) by OLS. This has both conceptual and practical issues. Conceptually, (12) is not a coherent causal model: it provides multiple, mutually inconsistent, expressions for the outcome, and the Stable Unit Treatment Value Assumption (SUTVA) is clearly violated. The causal effects of an extra election victory are also expected to vary substantially with s_j, limiting external validity of the OLS coefficient. For instance, de Benedictis-Kessner and Warshaw (2020) find the effect of an extra Democrat to be much larger in small legislatures. This issue would be avoided by an IV specification with the share of Democrats in the legislature as the treatment. Practically, in finite samples where the bandwidth is non-negligible, z_j may be non-trivially correlated with the treatment statuses of other subunits (both narrow and others). This potentially results in a form of asymptotic bias that doesn't arise in IV analyses.

4 Extensions: Spillovers and Temporal Aggregation

So far, we have considered settings in which outcomes are measured at a higher level of cross-sectional (e.g., spatial) aggregation than the RD events. In this section, we discuss two related settings to which our RDA ideas extend. Section 4.1 considers estimation of spillovers in RD settings, while Section 4.2 focuses on aggregating RD events over time. We describe how the two estimators from Section 2 apply in those settings and compare them with current practice.

4.1 Spillovers from RD Events

A number of studies consider spillovers from discontinuity events. Instead of units and subunits, we now have outcome units i and intervention units j.[15] We first discuss how our framework and solutions extend

[15]This terminology is common when analyzing experiments on bipartite networks (e.g., Zigler and Papadogeorgou (2021)).

to this setting, and then compare it with the current empirical practice.

In spillover cases, the endogenous variable X_i is often the number or fraction of treated neighbors, corresponding to the exogenous ("contextual") peer effects model. Thus, it still follows equation (2), with $\mathcal{J}_i$, reinterpreted as the set of i's peers. Alternatively, the endogenous linear-in-means peer effects model is sometimes used, in which X_i is the average outcome of the peers (Manski, 1993). In either case, the instrument is still defined by (4), with $\mathcal{C}_i$ reinterpreted as the set of i's peers with narrow elections.

Our upper-level solution applies with no change, while the lower-level solution can now take two forms. The stacked equation is now defined at the bilateral level: on a sample of *pairs* $\{(i,j)\colon j \in \mathcal{C}_i\}$, estimate

$$Y_i = \beta X_i + \gamma' \tilde{W}_i + \lambda' q_j + \text{error}_{ij}, \tag{13}$$

with z_j as the IV for X_i and s_j as weights. This is still a standard fuzzy RD specification with a continuous treatment, albeit with both i and j potentially observed multiple times. Absent additional i-level controls, i.e. if $\tilde{W}_i = 0$, one can also obtain the same estimate $\hat{\beta}$ by collapsing (13) to the level of intervention units, i.e. to the sample of narrow elections j, by averaging across outcome units that are j's neighbors: i.e., one can estimate

$$\bar{Y}_j = \beta \bar{X}_j + \lambda' q_j + \text{error}_j, \tag{14}$$

where $\bar{Y}_j = \frac{1}{N_j}\sum_{i\in\mathcal{I}_j} Y_i$ and $\bar{X}_j = \frac{1}{N_j}\sum_{i\in\mathcal{I}_j} X_i$ for $\mathcal{I}_j = \{i\colon j \in \mathcal{J}_i\}$ and $N_j = |\mathcal{I}_j|$, instrumenting $\bar{X}_j$ with z_j and using $s_j N_j$ as weights.[16] Versions of all three approaches have been used previously; yet, like in Sections 3.2 and 3.3, none of them exactly coincides with our proposal.

Surprisingly, we have found only one paper that performs estimation at the i level (Mora-Garćia and Rau, 2023). This stands in contrast not only with the many papers using the upper-level specification in the aggregation setting of Section 3.2, but more importantly with how peer effects are usually estimated in economics, via either exogenous or endogenous peer effects models (Manski, 1993). Ultimately, i-level specifications are, in our view, the most natural and coherent way to study i-level outcomes. Similar to papers in Section 3.2, Mora-Garćia and Rau (2023) do not include the third RDA control of equation (7) and only include the second RDA control as a robustness check.[17]

All other papers use specifications at the bilateral level or averaged by intervention unit j, and they are subject to the issues raised in Section 3.3. The first issue involves unnecessarily restricting the sample. We find this in Dahl et al. (2014) who entirely drop all units with more than one neighbor near the cutoff. The second issue is that most papers consider reduced-form OLS regressions on z_j, rather than introducing the treatment at the level of outcome units (or perhaps as in (14)). The only exception we are aware of is Dube et al. (2019) who estimate a bilateral IV specification similar to what we propose, with the average status of the peers as the endogenous variable. Other papers — e.g., Baskaran and Hessami (2018) and Dechezleprêtre et al. (2023) at the bilateral level and Santoleri et al. (2022) and Bhalotra et al. (2018) at the j level — are subject to the conceptual and practical concerns of reduced-form specifications discussed in Section 3.3.[18]

[16] Note that $\bar{X}_j$ is a convoluted treatment: it is the average treatment X_i of j's neighbors where X_i itself is an average or sum of its neighbors' RD shocks. The first-stage captures that, when j is treated, the average treatment exposure X_i of j's neighbors mechanically goes up.

[17] In addition, they limit the sample of outcome units to those outside the bandwidth and, in the main analysis, define the treatment by averaging only across the peers within the bandwidth. Our theory would apply even with the full sample and a more natural definition of the treatment based on all peers.

[18] These concerns also apply to papers that use IV estimation with a j-level treatment, rather than our proposed i-level treatment. Isen (2014), for instance, estimates the effect of Y_j on the average outcome of j's neighbors. This is in contract to a conventional endogenous peer effects model that would estimate the effect of the average outcome of i's neighbors on Y_i. See Clark (2009) and Dahl et al. (2014) for related examples.

4.2 RD Aggregation over Time

Another setting in which RDA ideas can be useful is when multiple RD events are aggregated over time. For instance, one may specify a state-level treatment as the share of years in the last decade when a Democratic governor is in power (in a cross-section of states or a panel with a rolling window). Our upper-level solution would entail instrumenting this treatment with the share of years in which the current governor is a Democrat who won in a narrow election, and controlling for the share of years determined by narrow elections, along with the other RDA controls. In turn, the lower-level solution would involve stacking all narrow elections of the previous decade, repeating the same outcome and aggregated treatment, and including standard RD controls. An unusual feature of the dynamic setting is that earlier elections can influence the outcomes of later elections, for instance via incumbency advantage (Lee, 2008). Our solutions allow for this, under a standard exclusion restriction.

A more complex version of this setting is when elections are not regular and the previous election's result can affect whether election is held in a future year. This is the case in the literature that studies the effects of public school funding on, e.g., house prices and educational outcomes. This literature uses RD variation from local referenda for school district bonds that authorize capital spending (see Jackson and Mackevicius (2024) for a review). If the bond is approved in one year, it is less likely that another referendum will be held the following year. In our lower-level solution, there will be endogenous selection of the set of narrow referenda included in the stacked sample. Still, nearly random outcomes of those referenda should make the estimator valid under an appropriate exclusion restriction (that precludes, for instance, a direct effect of holding a referendum on the outcome through an increase in media attention). Similarly, in our upper-level solution, RDA controls, such as the share of years with narrow referenda, may now resemble "bad controls," as they can be affected by the results of earlier referenda while also related to the error term. Yet, it is nearly by chance whether a narrow election results in the bond passing or not, conditional on prior referendum margins. Based on Proposition 1, we cautiously conjecture that the upper-level solutions also continues to be valid.

The current practice has followed a different approach, proposed by Cellini et al. (2010) in the school bond context. They offer several estimators; the "one-step" dynamic RD specification most frequently adopted in later studies is as follows:

$$Y_{it} = \sum_{h \in \mathcal{H}} \left(\beta_h x_{i,t-h} + \phi_h M_{i,t-h} + P\left(r_{i,t-h}; \theta_h\right) \right) + \tilde{\gamma}' \tilde{W}_{it} + \text{error}_{it}. \tag{15}$$

Here the set of β_h coefficients for a set $\mathcal{H}$ of lags ($h \geq 0$) and possibly leads ($h < 0$) aims to capture the dynamic effects of school spending $x_{i,t-h}$. The spending h periods ago is instrumented with the indicator of the bond approved in a referendum, $z_{i,t-h}$ (and reduced-form specifications are also considered). The set of controls includes the indicators $M_{i,t-h}$ that there was a referendum in year $t-h$ and global polynomials $P(r_{i,t-h}; \theta_h)$ in the vote margin $r_{i,t-h}$ with coefficients θ_h (where the vote margin is normalized to zero in years without an election), as well as additional controls $\tilde{W}_{it}$. Cellini et al. (2010) point out that their specification does not allow earlier elections to affect the occurrence of future elections or the future vote margin: if that happened, $M_{i,t-h}$ and $P(r_{i,t-h}; \theta_h)$ would be bad controls biasing the coefficients of interest β_h.[19]

[19] They also propose a "recursive" estimator which allows for some of these endogenous responses, although Hsu and Shen (2021) clarified that this is only true with homogeneous effects. The recursive estimator has not been popular in later work, perhaps because it tends to be noisier.

We view our RDA solutions as complementary to the dynamic RD specification. The key advantage of the Cellini et al. (2010) specification is that it allows to trace the dynamics of the effects (and also do placebo testing by looking at the lead coefficients). We note, however, that the effect of cumulative spending is also of interest: for instance, Biasi et al. (2024, Table 4), and Baron (2022, Tables 4–5) similarly report the average effect of the bond approval across different horizons. The key disadvantage of specification (15) is the potential bias from bad controls, which may not arise with RDA solutions, especially our lower-level specification. Additionally, (15) is based on global polynomial estimation, which is generally avoided in more recent RD analyses for being noisy and sensitive to the order of the polynomial (Gelman and Imbens, 2019), while the RDA solutions follow the prevalent local linear estimation approach.[20]

5 Application: The Impact of Unions on US Inequality

We now apply the RDA identification strategy to estimate the effect of unions on wage inequality within state-by-industry cells in the US. Section 5.1 introduces the empirical setting and adapts the RDA design to this application. Section 5.2 describes the data construction and presents summary statistics. Section 5.3 validates our research design in a series of balance tests. Section 5.4 presents the empirical results, and Section 5.5 discusses what these results mean for the contribution of declining unions to growing inequality in the US.

5.1 Empirical Setting and Specifications

Trade unions are significant labor market institutions in all Western countries, including the US. Over the last 100 years, union density in the US has followed an inverse U-shaped pattern: there was a sharp increase in unionization during the Roosevelt era (1930–45), a relative steady state of high unionization rates from 1945 to 1960, and a continuous decline since the 1960s (Farber et al., 2021, Figure 1). The strong negative correlation between this trend and the U-shaped pattern of U.S. income inequality (e.g., Piketty et al., 2018) contributes to the widespread perception of a relationship between the two phenomena.

Recent studies have explored the relationship between unionization and inequality (Callaway and Collins, 2018; Collins and Niemesh, 2019; Fortin et al., 2021; Farber et al., 2021). Although these studies report a significant negative association between unions and inequality in a cross-section or a panel, a causal interpretation of those results requires strong assumptions, as illustrated by the quote in the Introduction. Notably, these studies have not adopted the RD approach that is common in the analyses of the union effects on establishment-level outcomes (DiNardo and Lee, 2004; Sojourner et al., 2015; Frandsen, 2021).

Unionization in the US occurs at the level of a bargaining unit, defined as "a group of two or more employees who share community of interest and may reasonably be grouped together for purposes of collective bargaining" (National Labor Relations Board, 1997, p.12). In most cases, a bargaining unit corresponds to a single establishment, representing on average close to half of the establishment's workforce (Frandsen, 2021).[21]

[20] A local polynomial version of the IV specification (15) would replace $M_{i,t-h}$ with a dummy of holding a close referendum at $t-h$, replace $P\left(r_{i,t-h};\theta_h\right)$ with a local polynomial (filled in with a zero if there was no referendum at all or the referendum was not close), and replace $z_{i,t-h}$ with dummies of the bond approved in a close referendum as the set of instruments. We are not aware of any paper that used this estimation procedure or studied its properties formally.

[21] To be precise, Frandsen (2021, Table 1) reports an average of 93 votes and 254 total employees in his sample covering most unionization elections between 1980–2009. Using data from all unionization elections in this period, we found an average turnout rate of 85% (as a fraction of eligible voters), suggesting a typical bargaining unit size of approximately 109 employees (93/0.85). This indicates that bargaining units represented about 43% (109/254) of the total workforce in establishments holding elections.

The election typically offers two choices: for or against union formation. To form a union, a majority of 50% plus one out of the cast votes is necessary, lending itself to an RD design.

We exploit the discontinuities from narrow union elections to estimate union effects at the more aggregate level of state-industry cells. Specifically, we define the outcome ΔY_{sit} as the change in an economic outcome (e.g., a measure of wage inequality) in state s and industry i during decade t,[22] and the endogenous treatment NewUnions_{sit} as the share of newly unionized workers. The latter is measured as the ratio of the number of workers in the region-industry cell who join unions through unionization elections (both narrow and not) during the decade and the workforce in the cell at the beginning of the decade.

Looking at outcomes more aggregated than establishments has two advantages. First, it allows us to study the effects on outcomes, such as inequality, that are typically defined at more aggregated levels. Second, even for outcomes—such us average wages—that can be measured by establishment, our estimates capture the effects due to the spillovers of unionization on other establishments within the same cell. Establishments in the same state and industry often compete for the same workers and therefore may adjust wages or even enter or exit the market in response to another establishment's unionization. Defining cells as a combination of state and industry follows the approach of Fortin et al. (2021, 2022). Aggregating the analysis to purely geographic units — either states (Farber et al., 2021) or economic areas similar to commuting zones (such as Census State Economic Areas in Collins and Niemesh (2019)) — could help capture cross-industry spillovers, but such aggregation is not feasible for our study. State-level aggregation would substantially reduce statistical power, as typical cells would contain many close elections, and the law of large numbers would eliminate most identifying variation in our RDA approaches. Meanwhile, data limitations do not allow us to achieve spatial resolution smaller than states.[23]

Applying our upper-level IV solution from Section 2.2 to this setting, we estimate the following structural equation and first-stage:

$$\Delta Y_{sit} = \beta \text{NewUnions}_{sit} + \lambda' Q_{sit} + FE_{st} + FE_{it} + \text{error}_{sit}, \tag{16}$$

$$\text{NewUnions}_{sit} = \pi Z_{sit} + \rho' Q_{sit} + FE_{st} + FE_{it} + \text{error}_{sit}. \tag{17}$$

Here the instrument Z_{sit} is the employment share of the bargaining units unionized in a narrow election in the state-industry workforce. Formally, $Z_{sit} = \sum_{j \in \mathcal{C}_{sit}} s_j z_j$ where $\mathcal{C}_{sit}$ is the set of close elections in the si cell during decade t, z_j is the dummy that the vote share for the union strictly exceeds 50% in workplace j, and s_j is the number of eligible voters in election j as a fraction of the cell workforce at the beginning of the decade.[24] The covariates Q_{sit} consist of the three RDA controls defined by (5); in this context they represent the share of workforce involved in narrow elections ($\sum_{j \in \mathcal{C}_{sit}} s_j$), a rescaled average of the running variable r_k, defined as the union vote share minus 0.5, in closed elections ($\sum_{j \in \mathcal{C}_{sit}} s_j r_j$), and a rescaled average of the running variable in close wins ($\sum_{j \in \mathcal{C}_{sit}} s_j r_j^+$). Finally, FE_{st} and FE_{it} are state-by-decade and industry-by-decade fixed effects, respectively, included solely to increase estimation power, as the RDA design alone

[22] We use differenced outcome variables both to increase estimation power and as a way to reduce potential bias from manipulated unionization election results, as suggested by Frandsen (2017, 2021). Additionally, using differenced outcome variables aligns with our treatment—the rate of new unionization—representing most of the *change* in the share of unionized workers.

[23] For metropolitan areas, Census public files report geographic units less aggregated than states. However, both the number of these units and their boundaries change each decade, hindering the construction of panel data and the use of differenced outcomes.

[24] We distinguish between winning and obtaining a majority of the votes, as in a very small fraction of cases the final result was revised due to an appeal process. Our setting is "fuzzy" in this sense, and we handle it similarly to the standard fuzzy RD design: our instrument is defined based on dummies for crossing the 50% cutoff, while the treatment is based on dummies for union victory.

should be sufficient to obtain causal estimates.[25] We also explore specifications with different sets of fixed effects. To obtain more efficient and economically meaningful estimates, we weight each observation by the cell employment at the beginning of the decade.

We also apply the stacking approach of Section 2.3, which involves estimating the following structural equation and first-stage in the sample of narrow unionization elections j:

$$\Delta Y_{s(j)i(j)t(j)} = \beta \text{NewUnions}_{s(j)i(j)t(j)} + \lambda' q_j + FE_{s(j)t(j)} + FE_{i(j)t(j)} + \text{error}_j, \tag{18}$$

$$\text{NewUnions}_{s(j)i(j)t(j)} = \pi z_j + \rho' q_j + FE_{s(j)t(j)} + FE_{i(j)t(j)} + \text{error}_j. \tag{19}$$

where $s(j), i(j), t(j)$ respectively indicate the state, industry, and decade in which election j took place and q_j consist of the standard RD controls: a constant, the vote margin, and that margin interacted with the union getting a majority. The structural equation is identical to the upper level specification, except for the controls and the error term. Per Proposition 1, we weight observations by the number of eligible workers in each workplace j.

5.2 Data

We collect data on unionization and labor market outcomes by state, industry, and decade over five decades, from 1960s to 2000s.[26] The data on establishment-level union elections is from the National Labor Review Board (NLRB), maintained by Henry Farber and John-Paul Ferguson. For each election, we observe the state, industry, and year, as well as the total number of eligible workers and the votes cast for and against. We construct the treatment variable NewUnions_{sit} using the results of all elections. The instrument, in turn, aggregates only the elections that satisfy three criteria. First, it only includes close elections, defined by the vote share for the union of 50 ± 10 percent (the smallest bandwidth employed in recently published papers that employ the RD strategy in the context of US unions; Knepper (2020); Frandsen (2021)).[27] Second, we exclude tied elections and elections with a 1-vote margin, as there is evidence of manipulation in such elections (Frandsen, 2021). Finally, we drop elections with less than 20 votes, as is customary in the unionization literature (DiNardo and Lee, 2004; Lee and Mas, 2012; Frandsen, 2021).[28]

Our wage inequality measures and other wage-related outcomes by state, industry, and year are based on the decennial population census samples from 1960 to 2000, supplemented by the American Community Survey for 2009–2013 that represents 2010. With some variation over the years, these samples represent roughly 5% of the US population. We include observations from all 50 states, as well as from the District

[25] We do not include state-by-industry fixed effects both because the outcome is already measured in differences and because including such fixed effects would result in a substantial reduction of the sample: our industry definitions change after 2000, leading to only one post-2000 state-by-industry observation (see Section 5.2).

[26] This period was chosen based on the availability of unionization data, which is accessible from 1961 onwards and contains industry variables until the beginning of 2009.

[27] The 10p.p. bandwidth also closely aligns with the optimal bandwidth for local linear estimation. Applying the Calonico et al. (2014) procedure to our lower-level specification across our main five inequality outcomes, we find optimal bandwidths ranging from 9.5% to 13.7%, with an average of 10.9%. We also check robustness to using a 15% bandwidth which is also used by Knepper (2020) and Frandsen (2021).

[28] During the period of our analysis, there are two types of unionization elections: certification elections, where workers attempt to form a new union, and decertification elections, where workers or employers try to dissolve an existing union. The rules for both types of elections are identical, as are the consequences: if the union receives a strict majority of votes, it will be certified or re-certified; if not, it will be decertified. The vast majority of elections are certification elections, accounting for 88.2% of elections and covering 90.2% of the workers. We do not distinguish between these two types of elections in our analysis; e.g., $NewUnions_{sit}$ counts workers in both newly certified and re-certified unions. Our results are robust to restricting the set of elections used for the treatment and instrument construction to certification elections only.

of Columbia, defined as an additional state. We only include workers with a strong attachment to the labor market, defined as working all 52 weeks during the year and at least 20 hours per week on average.[29] Two of our main inequality outcomes are computed using total labor income: the Gini index and the top-10% income share. For the other measures, we calculate the average annual hourly pre-tax wage of each worker by dividing the total annual wage income by the product of the number of weeks employed and the average hours worked per week last year. We then compute the college wage premium, the log 90-10 wage ratio, and the variance of log wages, again by state-industry cell. The construction of the inequality measures is detailed in the Data Appendix, following Autor et al. (2008) and Farber et al. (2021) with minor modifications. For the inequality measures to be meaningful, we restrict the sample to include only cells with more than 1,000 workers after applying census projection weights.[30]

We additionally use the Current Population Survey (CPS) Monthly and Annual Social and Economic Supplement samples. We measure union density in each cell and year based on the self-reported union coverage question. We also construct two measures of worker benefits: indicators of a pension plan and employer-sponsored health insurance coverage. These questions are available prior from 1977–2014; to increase sample size we combine CPS data from several years to represent each decennial census year, as detailed in the Data Appendix. The cleaning process of the CPS data is otherwise the same as for the census data.

A challenge in the data construction process was to harmonize industry definitions over a long period and between different data sources. The NLRB data is based on the SIC and NAICS schemes, depending on the year, while the population census, ACS, and CPS use a different classification. We created two distinct coding schemes. Until 2000, our industry coding is based on the initial two digits of the SIC87 codes; from 2000 onwards, it relies on the initial three digits of the NAICS97 industry classification. In both cases, we merged a few categories, resulting in 70 industries categories pre-2000 and 73 (different) categories post-2000 that cover all private sectors.[31] To calculate differenced outcomes for each decade, we estimate all measures for the year 2000 based on both schemes.[32] In total, we were able to match 93% of unionization elections to our industry codes and 99% of individual census observations.

Table 1 presents summary statistics for the key variables at the state-industry-census year level, both in levels and first-differences. Panel A includes outcomes, while Panel B details other cell-level statistics. One can see, in particular, that all within-cell inequality measures show an increasing trend, mirroring national inequality trends. Panel C reports the endogenous variable, the instrument, and the RDA controls described above, as well as the average numbers of all and narrow unionization elections. In an average decade, there are 55 union elections per cell resulting in 2.5% of the workforce joining the union. An average of 8 elections per cell, involving 0.6% of the workforce, results in a narrow win that enters our instrument construction. Appendix Table A2 reports summary statistics at the election level, showing in particular that the number of elections has seen a continuous decline since 1970s (which matches the negative trend in union membership reported in Table 1, Panel A, with the CPS data since 1980s).

While we only look at inequality within state-industry cells, this type of inequality is central to both the

[29] We exclude from the sample workers who worked fewer than 52 weeks, as they have a higher likelihood of having switched industries during the past year, in which case we could attribute their annual income to the wrong industry.

[30] This entails dropping 24.9% of the cells, but only 0.8% of total employment and 1.3% of close unionization elections. We verify there are no remaining cells with fewer than 10 raw observations. For measures based on smaller samples, such as for the average college wage, we further exclude cells with fewer than 10 relevant observations at the beginning or at the end of the decade.

[31] The NLRB data do not cover public sector unionization and, thus, those industries are excluded from our analysis. Our classification is more detailed than 10–11 industries used by Fortin et al. (2021, 2022).

[32] Since the 2000 census industry question was based on the NAICS industry scheme, we relied on the IPUMS crosswalk to match observations to the SIC-based census industries. While this could introduce noise in the 1990–2000 differences, we find it reassuring that employment changes do not exhibit substantially higher volatility than in the decade before or after.

Table 1: Summary Statistics

Variable	Mean	Mean(Δ)	Data Source	Time Range
Panel A: Outcomes				
Log avg hourly wage	2.825	0.043	Census	1960-2010
	(0.284)	(0.119)		
Log college premium	0.505	0.049	Census	1960-2010
	(0.164)	(0.106)		
Log 90/10 ratio	1.314	0.070	Census	1960-2010
	(0.216)	(0.109)		
Gini index	0.347	0.023	Census	1960-2010
	(0.070)	(0.032)		
Top 10% share	0.282	0.017	Census	1960-2010
	(0.058)	(0.032)		
Var(log hourly wage)	0.310	0.032	Census	1960-2010
	(0.097)	(0.048)		
Share of union members	0.102	-0.040	CPS	1980-2010
	(0.125)	(0.078)		
Pension coverage share	0.635	-0.009	CPS	1980-2010
	(0.188)	(0.119)		
Employer-sponsored insurance share	0.678	-0.062	CPS	1980-2010
	(0.164)	(0.096)		
Panel B: Other Cell-Level Measures				
Log employment	10.865	0.170	Census	1960-2010
	(1.254)	(0.309)		
Share with college degree	0.241	0.041	Census	1960-2010
	(0.177)	(0.041)		
Share white	0.835	-0.025	Census	1960-2010
	(0.119)	(0.042)		
Share male	0.586	-0.017	Census	1960-2010
	(0.223)	(0.041)		
Panel C: RDA Variables (bandwidth = 10%)				
$100 \times \text{NewUnions}_{sit}$ (Share newly unionized)	2.475		NLRB	1960-2010
	(5.035)			
$100 \times Z_{sit}$ (Share newly unionized in close elections)	0.634		NLRB	1960-2010
	(1.591)			
$100 \times \sum_{j \in \mathcal{C}_{sit}} s_j$ (Share involved in close elections)	1.530		NLRB	1960-2010
	(3.342)			
$100 \times \sum_{j \in \mathcal{C}_{sit}} s_j r_j$ (Agg. vote share in close elections)	-0.018		NLRB	1960-2010
	(0.118)			
$100 \times \sum_{j \in \mathcal{C}_{sit}} s_j r_j^+$ (Agg. vote share in close wins)	0.033		NLRB	1960-2010
	(0.083)			
# of unionization elections	55.2		NLRB	1960-2010
	(83.7)			
# of close unionization elections	8.2		NLRB	1960-2010
	(14.7)			

Notes: The first column displays the weighted averages and (in parentheses) standard deviations of the end-of-decade levels of the variables, while the second column reports these statistics for first differences over the decade. The statistics are weighted by employment at the beginning of the decade. Observations are state-industry cells based on SIC industry codes pre-2000 and NAICS codes thereafter. Panel A summarizes the outcome variables used in the analysis. Panel B provides additional covariates. Panel C presents the RDA variables (scaled by a factor of 100) along with the average numbers of all and close unionization elections within each cell.
Sources: Decennial Census, CPS Annual and Monthly Social and Economic Supplement, NLRB data, and authors' calculations.

level and growth in overall inequality. To show this, Appendix Table A3 decomposes the overall variance of log wages into within- and between-cell components in each census year, broadly following Helpman et al. (2016). Throughout the period, the within component of the variance is the dominant one, accounting for 73–80% of total inequality. Moreover, the within component experienced a continuous rise, contributing significantly to the overall increase in inequality. In contrast, the between-group component remained relatively stable, decreasing in some periods and only experiencing a notable rise from 2000 to 2010.[33]

5.3 Balance

We now perform balance tests to illustrate how the RDA instrument and controls help isolate idiosyncratic variation in the treatment. Column 1 of Table 2 regresses the share of all newly unionized workers (i.e., our treatment of interest) on pre-determined covariates and lagged inequality measures from Table 1 (excluding those which are not available for all decades). The treatment is correlated with many observables, as evidenced by the partial R^2 of 13.5% and a decisively rejecting F-test, which suggests that the treatment is also likely correlated with unobservables. In column 2 we estimate the same regression for our instrument: the share of workers who join a union through close elections; the conclusion is unchanged. In column 3, the inclusion of the first RDA control that represents the share of workers involved in narrow union elections vastly changes the statistics. The partial R^2 decreases massively to 0.1%, the F-statistic of the covariates' joint significance shrinks from 185.6 to 3.5, and most coefficients become orders of magnitude smaller. Yet, covariates are still jointly significant (with a p-value of 0.01%). When we finally add the two other RDA controls to the regression in column 4, the partial R^2 becomes essentially zero and, while one out of ten covariates is individually significant, jointly they are not (p-value of 12.3%). The lack of correlation between the residual variation in our instrument and the observables speaks in favor of the validity of our IV strategy. In Appendix Table A4 we replicate this analysis with a 15% bandwidth defining close elections, and the same patterns emerge.

Appendix Table A5 performs a related pre-trend and placebo analysis, using predetermined covariates and lagged outcomes, both in levels and differences, as dependent variables when estimating the reduced-form of our upper-level IV specification:

$$Y_{sit} = \rho Z_{sit} + \lambda' Q_{sit} + FE_{st} + FE_{it} + \text{error}_{sit}. \tag{20}$$

None of the 15 coefficients are significant at the 5% level using either conventional standard errors or robust bias-corrected confidence intervals of Calonico et al. (2014).

5.4 Results and Mechanisms

Table 3 examines the impact of unionization on five key measures of income inequality: the college premium, the 90-10 wage ratio, the Gini coefficient, the top-10% income share, and the variance of log hourly wages. Panel A shows the estimates using the upper-level approach, while Panel B shows the coefficients obtained

[33] A concern in this analysis is that some of the changes in the variance components are due to the changes in the definitions of industries in the year 2000 from SIC based to NAICS based. To address this, we conducted an identical analysis on with annual CPS samples in Appendix Figure A1. It indicates no discernible changes in the within and between components between 2002 and 2003, the year of the CPS industry scheme change, providing strong evidence that this transition does not drive the trends in the components. This analysis also replicates the findings from Appendix Table A3 using different data, confirming the key role of within-cell inequality.

Table 2: Instrument Balance

	NewUnions	Z		
	(1)	(2)	(3)	(4)
Log avg hourly wage$_{t-1}$	1.056	0.153	-0.100	0.009
	(0.517)	(0.095)	(0.055)	(0.024)
Log college premium$_{t-1}$	4.452	1.812	0.083	-0.006
	(0.405)	(0.133)	(0.079)	(0.034)
Log 90/10 ratio$_{t-1}$	3.895	0.889	-0.010	0.044
	(0.630)	(0.194)	(0.108)	(0.049)
Gini index$_{t-1}$	-58.022	-14.715	-0.876	0.060
	(6.424)	(1.434)	(0.836)	(0.358)
Top 10% share$_{t-1}$	40.646	9.672	0.641	-0.005
	(6.474)	(1.344)	(0.789)	(0.340)
Var(log hourly wage)$_{t-1}$	-7.027	-1.805	0.012	-0.181
	(1.486)	(0.287)	(0.170)	(0.077)
Log employment$_{t-1}$	-0.158	-0.060	0.003	-0.004
	(0.051)	(0.015)	(0.008)	(0.003)
Share with college degree$_{t-1}$	-4.510	-1.035	0.129	0.067
	(0.609)	(0.126)	(0.076)	(0.040)
Share white$_{t-1}$	-0.527	-0.096	-0.257	-0.024
	(0.561)	(0.174)	(0.094)	(0.040)
Share male$_{t-1}$	-0.869	-0.377	-0.015	0.027
	(0.516)	(0.126)	(0.063)	(0.025)
# significant coefficients	9	9	2	1
Partial R^2	0.1353	0.0959	0.0012	0.0001
Partial F-statistic	292.7	185.6	3.5	1.5
Partial F-test, p-value	0.0000	0.0000	0.0001	0.1234
RDA controls	None	None	$\sum_{j \in \mathcal{C}_{sit}} s_j$	All
Observations	12,799	12,799	12,799	12,799

Notes: Balance tests based on OLS estimates weighted by employment at the beginning of the decade. The outcome variable in the first column is the share of all newly unionized workers, while the next three columns focus on shares joining through close elections, with a bandwidth of 10% applied to define close elections. Column 3 additionally incorporates the first RDA control, $\sum_{j \in \mathcal{C}_{sit}} s_j$, and column 4 adds the remaining RDA variables, $\sum_{j \in \mathcal{C}_{sit}} s_j r_j$ and $\sum_{j \in \mathcal{C}_{sit}} s_j r_j^+$. Heteroskedasticity-robust standard errors are shown in parentheses.
Sources: Decennial Census, NLRB data, and authors' calculations.

Table 3: The Impact of New Unionization on 10-Year Changes in Inequality

	ΔLog college premium (1)	ΔLog 90/10 (2)	ΔGini coeff. (3)	ΔTop 10% share (4)	ΔVar(log wage) (5)
Panel A: Upper-Level Estimator					
Share newly unionized	-0.314	-0.459	-0.176	-0.143	-0.255
	(0.285)	(0.225)	(0.061)	(0.055)	(0.084)
	[-0.921,0.416]	[-0.933,0.122]	[-0.325,-0.035]	[-0.279,-0.020]	[-0.463,-0.054]
Mean outcome	0.515	1.318	0.348	0.282	0.311
Mean treatment	2.65%	2.76%	2.76%	2.76%	2.76%
RDA controls	✓	✓	✓	✓	✓
Industry-decade and state-decade FE	✓	✓	✓	✓	✓
Observations	10,512	12,910	12,910	12,910	12,910
Panel B: Lower-Level Estimator					
Share newly unionized	-0.258	-0.266	-0.133	-0.115	-0.206
	(0.284)	(0.185)	(0.052)	(0.048)	(0.073)
	[-0.896,0.443]	[-0.735,0.140]	[-0.284,-0.035]	[-0.256,-0.026]	[-0.427,-0.075]
Mean outcome	0.492	1.189	0.305	0.251	0.254
Mean treatment	9.00%	9.54%	9.54%	9.54%	9.54%
RD controls	✓	✓	✓	✓	✓
Industry-decade and state-decade FE	✓	✓	✓	✓	✓
Observations	29,146	31,256	31,256	31,256	31,256

Notes: Panel A reports IV estimates of (16). Observations are defined as state-industry cells by decade (for 1960–2010). The endogenous variable is the share of workforce unionized through unionization elections, relative to the beginning-of-decade cell workforce. The instrument is the share of workforce unionized through close elections, defined by the 10% bandwidth for the vote share around the cutoff of 50%. RDA controls are as in equation (20), and state-year and industry-year FEs are included. See Data Appendix for the outcome definitions and details of the sample construction. Observations are weighted by employment at the beginning of the decade. Panel B shows the coefficients obtained from the lower-level IV specification (18), in the sample of close elections. The specification includes standard RD controls and is weighted by the number of eligible workers. In both panels, heteroskedasticity-robust standard errors are in parentheses, and robust bias-corrected confidence intervals of Calonico et al. (2014) are in brackets (in Panel A, constructed via the equivalent lower-level specification as in Proposition 1).
Sources: Decennial Census and NLRB unionization data; authors' calculations.

at the lower level. Along with each estimate, we report robust bias-corrected confidence intervals following Calonico et al. (2014), obtained using the `RDrobust` package in `R`.[34]

The results reveal a negative effect of unionization on all measures of inequality, with 95% confidence intervals rejecting zero effects for the last three inequality measures and with large but statistically insignificant negative effects for the college premium and the 90-10 wage ratio. Specifically, upper-level estimates suggest that an increase of 1 percentage point in new unionization reduces the college premium by 0.31 log-points, the 90-10 ratio by 0.46 log-points, the Gini coefficient by 0.018, the top-10 income share by 0.14 percentage points, and the variance of log wages by 0.0025 (which is around 1% of its mean). Lower-level estimates are moderately smaller in magnitude but show the same statistical significance. Figure 2 visualizes these results with standard RD plots for the first-stage and reduced-form lower-level specifications (excluding the fixed effects, to stay closer to the raw data).

Appendix Table A6 reports several robustness tests for these results, focusing on the upper-level analyses. First, we replace our baseline bandwidth of 10% with 15% bandwidth, a common *ad hoc* choice in the literature (e.g., Frandsen (2021); Knepper (2020)). Second, we exclude union decertification elections from the construction of both the treatment and the instrument (along with the RDA controls). Third, we replace state-year and industry-year fixed effects with separate state, industry, and year and FEs, or with no FEs at all. Finally, we perform estimation without initial employment weights. All results are similar to those from Table 3, with coefficients from the unweighted estimation generally slightly larger, possibly indicating a larger union effect on inequality in smaller cells.

While we specified the treatment variable to be the rate of *new* unionization, the results would also be similar if we considered changes in union density, i.e. in the overall share of unionized workers. The latter measure would take into account differential employment growth of unionized and non-unionized establishments, as well as establishment entry and exit. Unfortunately, union density is only observed in the CPS since 1977. Thus, in the spirit of two-sample IV (Angrist and Pischke, 2009, Ch.4.3), we estimate the effect of new unionization rates (from the NLRB data) on union density (from the CPS monthly samples) for the shorter period of 1980–2010.[35] The coefficient is reported in the first column of Table 4. This coefficient may be smaller than one due to various types of measurement error: both on the right-hand side (discrepancies arising from out-of-state workers, errors in coding the industry of unionization, or lower employment growth in unionized establishments) and on the left-hand side (misclassification of the union coverage variable in the CPS leading to a non-classical measurement error).[36] Conversely, if unionized workplaces have higher employment growth, the first-stage parameter might exceed one. In the data, we find estimates ranging from 0.63 to 1.27, suggesting that our main IV estimates are robust to using the change in the self-reported share of total unionized workers as the endogenous variable.

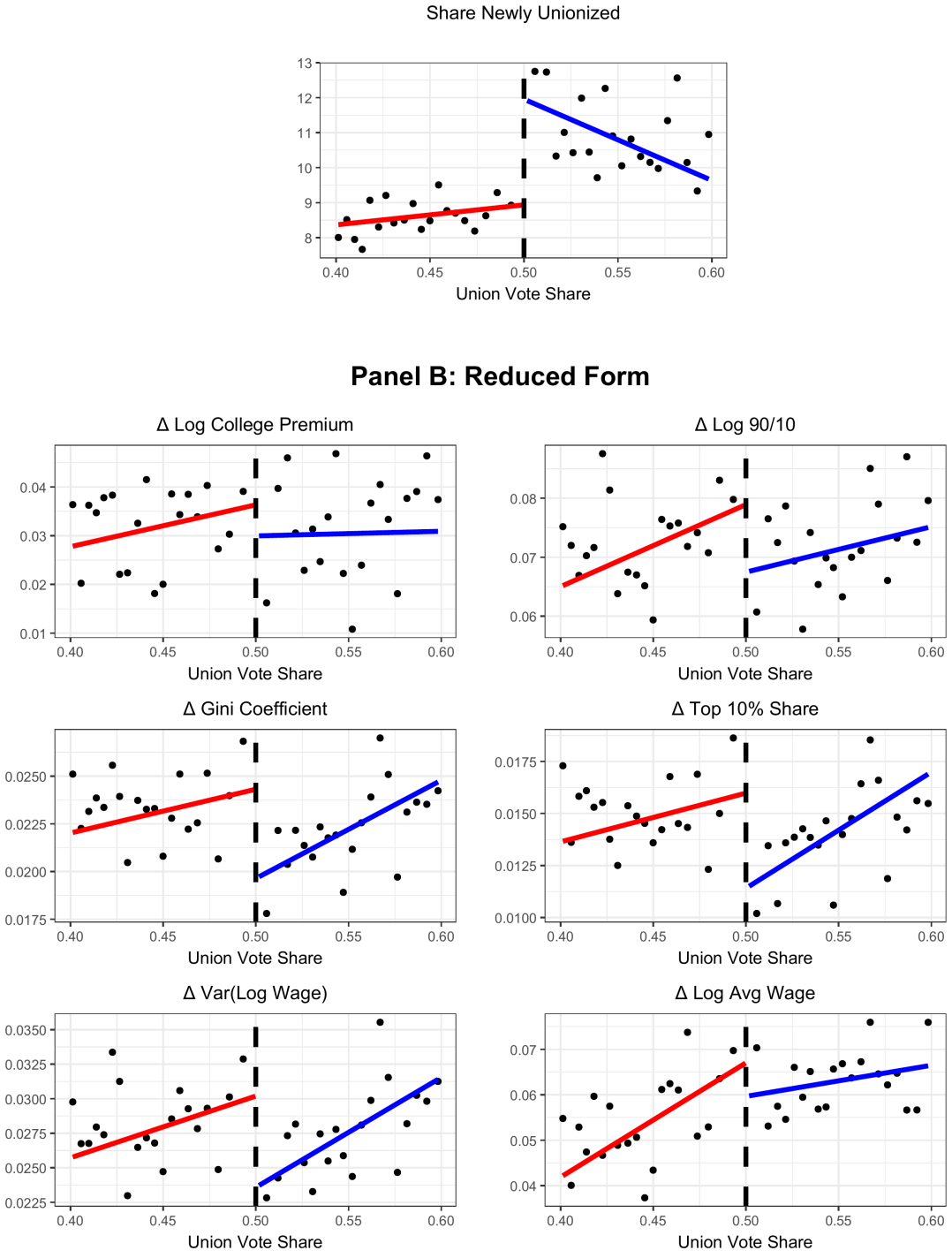
Table 5 investigates the sources of the inequality reduction due to unionization, finding that it originates primarily from the decline in earnings among high-earners, rather than an increase at the bottom of the earn-

[34] To obtain confidence intervals for our upper-level approach, we use the numerically equivalent lower-level specification, per Proposition 1. We do not use clustered inference (e.g., by state or by industry) for several reasons: theoretically, election shocks should be mutually independent; empirically, we confirm that clustered standard errors do not substantially exceed heteroskedasticity-robust ones; and pragmatically, the implementation of clustering in the `RDrobust` package does not allow elections on the two sides of the cutoff to be part of the same cluster.

[35] One may think of this analysis as a first-stage, except that this is an IV specification leveraging close elections for the instrument as before.

[36] Based on a 1977 validation survey, the estimate for the misclassification rate in the union variable is 2.7% (Card, 1996).

Figure 2: Lower-Level Specifications: RD Plots



Notes: The first plot visualizes the first-stage specification (19) with no fixed effects. The other plots visualize the reduced-form version of (18) with no fixed effects for the outcomes included in Table 3 and column 1 of Table 5. The sample, weights, and outcome definitions are the same as in Panel B of the relevant tables. The x-axis represents the union vote share in a close union election, while the y-axis shows the outcome in the corresponding region-industry cell and year. The binned scatterplot shows 20 points on each side of the cutoff, with each point representing the average values for elections within specific ranges of the running variable. These ranges were selected so that, on each side of the cutoff, each point represents observations with the same sum of weights. The lines correspond to the local linear estimates. "Δ" stands for the first difference over a decade.
Sources: Decennial Census and NLRB unionization data; authors' calculations.

Table 4: The Impact of New Unionization on 10-Year Changes in Union Density and Fringe Benefits

	Union density (1)	Pension coverage rate (2)	Health insurance coverage rate (3)
Panel A: Upper-Level Estimator			
Share newly unionized	1.274	1.482	0.063
	(0.551)	(0.780)	(0.466)
	[-0.043,2.490]	[-0.376,3.122]	[-1.010,0.986]
Mean outcome	0.105	0.522	0.680
Mean treatment	1.24%	1.49%	1.49%
RDA controls	✓	✓	✓
Industry-decade and state-decade FE	✓	✓	✓
Observations	7,846	6,630	6,675
Panel B: Lower-Level Estimator			
Share newly unionized	0.634	0.858	0.148
	(0.323)	(0.496)	(0.327)
	[-0.132,1.481]	[-0.285,2.090]	[-0.645,0.806]
Mean outcome	0.153	0.569	0.737
Mean treatment	3.90%	4.38%	4.43%
RDA controls	✓	✓	✓
Industry-decade and state-decade FE	✓	✓	✓
Observations	14,119	13,715	13,743

Notes: This table reports IV estimates from the same specifications as in Table 3 but for different outcomes: union density, pension coverage rate, and employment-sponsored insurance coverage rate. The outcomes are based on the CPS data and measured for the period from 1980–2010.
Sources: Decennial Census, CPS Annual and Monthly Social and Economic Supplement, NLRB unionization data; authors' calculations.

Table 5: The Impact of Unions on 10-Year Wage Changes

	All workers (1)	College (2)	High school (3)	P90 (4)	P50 (5)	P10 (6)	Mgmt (7)	Non-Mgmt (8)
Panel A: Upper-Level Estimator								
Share newly unionized	-0.348	-0.449	-0.105	-0.282	-0.129	0.176	-0.917	-0.181
	(0.150)	(0.349)	(0.142)	(0.180)	(0.145)	(0.166)	(0.399)	(0.134)
	[-0.762,-0.032]	[-1.221,0.468]	[-0.442,0.227]	[-0.735,0.138]	[-0.533,0.188]	[-0.283,0.497]	[-1.854,0.058]	[-0.561,0.091]
Mean outcome	2.832	0.545	-0.019	3.305	2.637	1.986	0.507	0.106
Mean treatment	2.76%	2.65%	2.76%	2.76%	2.76%	2.76%	2.71%	2.76%
RDA controls	✓	✓	✓	✓	✓	✓	✓	✓
Industry-decade and state-decade FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	12,910	10,541	12,873	12,910	12,910	12,910	10,638	12,910
Panel B: Lower-Level Estimator								
Share newly unionized	-0.303	-0.380	-0.119	-0.158	-0.173	0.108	-0.924	-0.130
	(0.128)	(0.351)	(0.121)	(0.151)	(0.127)	(0.136)	(0.386)	(0.111)
	[-0.725,-0.100]	[-1.237,0.433]	[-0.44,0.134]	[-0.584,0.146]	[-0.547,0.061]	[-0.241,0.395]	[-1.946,-0.127]	[-0.492,0.048]
Mean outcome	2.809	0.525	-0.010	3.253	2.658	2.064	0.468	0.057
Mean treatment	9.54%	9.00%	9.54%	9.54%	9.54%	9.54%	9.20%	9.54%
RDA controls	✓	✓	✓	✓	✓	✓	✓	✓
Industry-decade and state-decade FE	✓	✓	✓	✓	✓	✓	✓	✓
Observations	31,256	29,148	31,250	31,256	31,256	31,256	30,171	31,256

Notes: This table reports IV estimates from the same specifications as in Table 3 but for different outcomes. The outcome in column 1 is the change in log average wage in the sample of workers. In columns 2–3 and 7–8, the outcome is the change in the log average wage among specific groups of workers: those with exactly a college degree, exactly a high school degree, managerial occupations, and non-managerial occupations, respectively. State-industry cells with fewer than ten observations for the relevant group at the beginning or end of the decade are excluded. In columns 4–6, the outcome is the change in the weighted 90th, 50th, and 10th percentile of the distribution of log hourly wages within the cell. Observations are weighted by total cell-level employment at the beginning of the decade in all columns.
Sources: Decennial Census and NLRB unionization data; authors' calculations.

ings distribution. Column 1 reports a large negative and marginally significant effect on average wages.[37,38] The following columns of Table 5 measure the impacts of unionization on average wages of different worker groups and parts of the wage distribution. Columns 2 and 4 indicate a negative (albeit not statistically significant) impact of unions on wages among the higher-earning segments of the labor force: college graduates and 90th percentile earners, respectively; column 7 indicates a large and marginally significant negative effect on managers' earnings. In contrast, columns 3, 5, and 8 show negative but small and statistically insignificant wage impacts on the lower segments: high school graduates, 50th percentile earners, and non-managers, respectively. Column 6 finds a small and insignificant increase in wages at the 10th percentile of the distribution. By construction, the more negative effects at the top are consistent with our main finding of the negative impact of unions on inequality.[39]

The null, or even negative, effects on median and average wages, is puzzling and may raise questions about the motivations for unionization and the resistance of employers to unionization. We shed light on a possible explanation by considering the effects of unionization on fringe benefits—health insurance and pensions coverage—in columns 2 and 3 of Table 4. Since both the outcomes and the endogenous variable are measured in percentages, the coefficient is interpreted as the number of covered workers induced by one new union member. For health insurance, the effect is small and insignificant. However, for pension coverage, we estimate that each new union member corresponds to an increase of 1.48 pension holders (standard error of 0.78) using the upper-level instrument and 0.86 (SE of 0.50) when the lower level-instrument is employed. Although the coefficients are noisy, the fact that they are around one or even larger suggests a substantial spillover effect of unionization across establishments within state-industry cells, given that some of the unionized establishments had pension plans prior to unionization. This is consistent with the recent findings by Knepper (2020) that unionization of establishments within firms has significant impacts on employer pension contributions on other establishments of the same firm. Our analysis may also suggest spillovers to other firms within the same state and industry.

5.5 How Much Did Declining Unions Contribute to Growing Inequality?

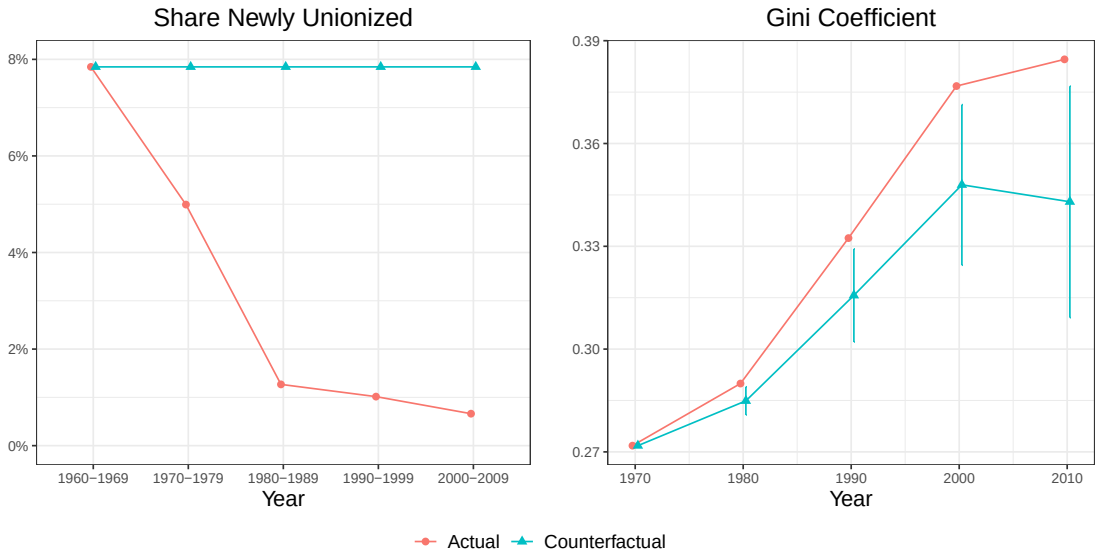
Rates of new unionization in the private sector in the US experienced a massive and continuous decline since the 1970s, as can be seen in the first panel of Figure 3. In the 1960s, the net flow of unionized workers through unionization elections was 7.8%. This is the result of the share of workers joining unions through unionization elections (8.3%), minus some workers leaving unions through union decertification elections.

[37] This large negative effect is unusual relative to the other estimates of a union wage premium at the worker or establishment level. Card (1996) and Farber et al. (2021, Figure 5) find a positive effect using a selection-on-observables design at the individual level. Staggered difference-in-differences estimates from De Chaisemartin and d'Haultfoeuille (2020) also at the individual level and RD estimates by DiNardo and Lee (2004) and Frandsen (2021) at the establishment level point out to a null or slightly negative effect. Our estimate may differ for several reasons. First, we incorporate spillover effects across establishments in the same state and industry. Second, our estimates incorporate spillovers on non-unionized employees, such as management (including even managers located at a different establishment), while worker-level estimates do not. Finally, like any RD-based estimate, we only use local variation from close elections.

[38] A caveat to this finding is that it may partially result from worker selection, as unionization has unequal effects on employment across worker groups with different earnings potential, as studied in Appendix Table A7. There is no clear negative effect on overall employment (Column 1), nor on the employment of high-school graduates and non-college workers more broadly (Columns 2–3). In contrast, employment of workers with exactly a college degree or those with college or more education declines substantially in response to unionization (Columns 4–5, marginally insignificant). This finding aligns with Frandsen (2021) who documents similar compositional effects at the establishment level, where high-paid workers tend to leave following unionization. Such composition effects can appear as a reduction in average wages. One may expect the composition effects to be smaller for columns 2–3 and 7–8 of Table 5 that investigate less heterogeneous worker populations.

[39] Appendix Table A8 conducts several robustness checks for these results. The overall conclusion of a negative effect on average wages driven by the high-earners is preserved, although the negative effect on the managerial wages tends to be weaker in the robustness analyses.

Figure 3: The Counterfactual with Unionization Rates Remaining at the 1960s Level



Notes: The first panel presents the actual and counterfactual rates of newly unionized workers for each decade between 1960 and 2010. The actual rate is calculated as the share of workers joining unions through unionization elections minus the share leaving due to union decertification elections. The counterfactual rate assumes that the rate remained constant at its 1960s level. The second panel presents the actual and counterfactual Gini coefficients of income inequality within state-industry cells, averaged across cells (with beginning-of-decade employment weights). The counterfactual is calculated by adding to the actual Gini the product of the unionization effect estimated in Table 3, column 3, and the cumulative shortfall in the rate of new unionization between the actual and counterfactual scenarios. Confidence intervals for the counterfactual Gini are based on the 95% robust bias-correct confidence intervals for the treatment effect.
Sources: Decennial Census, NLRB unionization data; authors' calculations.

The net flow declined consistently, reaching less than 1% in the 2000s.

Using the results in Table 3, we conduct a back-of-the-envelope calculation of how much of the growth in US inequality during the period of our analysis can be attributed to this sharp decline. To this end, we compute counterfactual inequality measures that may have occurred if new unionization rates had remained at their 1960s levels. Specifically, we add to each actual inequality measure the product of the relevant treatment effect and the cumulative shortfall in new unionization resulting from the sharp decline after the 1960s.

Like any back-of-the-envelope calculation, ours requires some caveats. The estimates from Table 3 capture a short-term effect and are identified using local variation from narrow elections. Here, we extrapolate these effects to the long term, assuming that the effect of new unionization during a decade on the level of inequality by the end of the decade equal the effect on inequality at later dates, too. Our analysis is also limited to assessing the effects of unions on inequality within state-industry cells. However, as shown above, within-cell inequality accounts for the majority of overall inequality both in levels and in changes.[40]

With these caveats, the second panel of Figure 3 presents the actual and counterfactual average within-cell Gini coefficients by census year. The gap between the actual and counterfactual inequality is small in 1980 and 1990 but becomes substantial by 2000 and 2010. Appendix Figure A2 reports similar results for the other inequality measures.

Table 6 summarizes all results of this back-of-the-envelope exercise showing a sizable contribution of

[40]Another limitation is that we assess only the effect of new unionization, not the full impact of changes in union density, which may also result from the closure or growth of unionized establishments. However, this concern is mitigated by our earlier finding that the effect of new unionization on union density is approximately equal to 1.

Table 6: Counterfactual Inequality Measures

Inequality measure	Actual			Counterfactual		Union contribution
	1970	2010	Change	2010	Change	as % of change
Log college premium	0.430	0.579	0.149	0.489	0.059	60.5%
				(0.062)		(41.8%)
Log 90/10 ratio	1.113	1.428	0.314	1.322	0.208	33.7%
				(0.058)		(18.3%)
Gini coefficient	0.272	0.385	0.113	0.346	0.074	34.5%
				(0.015)		(13.6%)
Top 10% share	0.226	0.310	0.084	0.278	0.052	38.3%
				(0.014)		(17.3%)
Var(Log wage)	0.220	0.362	0.143	0.310	0.091	36.4%
				(0.021)		(15.0%)

Notes: This table presents statistics regarding the five measures of wage inequality within state-industry cells used in Table 3. The 2010 counterfactual measures are calculated by adding to each actual inequality measure the product of the relevant treatment effect and the cumulative shortfall in new unionization resulting from the sharp decline since the 1960s. "Change" refers to the change from the actual measure in 1970 to the actual or counterfactual measure in 2010. The "Union Contribution" is calculated as one minus the counterfactual change in each measure divided by the actual change. The standard errors in parentheses are based on the standard errors for the corresponding estimates in Table 3.
Sources: Decennial Census, NLRB unionization data; authors' calculations.

declining unions to the growing within-cell inequality between 1970 and 2010. Unions account for 34-38% of the changes in the Gini coefficient, the 90/10 ratio, the top 10% share, and the variance of log wages. The contribution to the log college premium is even higher (at 61%) but imprecisely estimated. These estimates are moderately larger than the Farber et al. (2021, Table 1) estimates of the contribution of unions to within-state inequality in the US between 1968 and 2014.

6 Conclusion

We introduced a novel extension to the regression discontinuity design, termed Regression Discontinuity Aggregation (RDA), which aims at identifying local average causal effects in contexts where each unit may be exposed to multiple discontinuity events. The RDA approach is applicable when the outcome is defined at a higher level of aggregation (e.g., in space or time) than the discontinuity events or when spillovers of discontinuity events are of interest. Our theoretical analysis rests on the observation that an aggregation of RD events can be viewed as a special case of a shift-share variable. This connection allowed us to build two estimators for causal effects in RDA settings that build on local linear estimation in conventional sharp and fuzzy RD designs and inherit their bias-reduction properties.

To illustrate the usefulness of this approach, we applied the proposed estimation techniques to the effects of unionization on earnings inequality at the level of state-by-industry cells. We found that increased rates of new unionization significantly reduce inequality, primarily by lowering top-tier wages rather than raising lower-end wages. As a result, the decline of unions since 1970s can explain a substantial fraction of increased inequality. We additionally provide evidence that unionization leads to increased pension coverage, potentially in non-unionized establishments, too. Our findings provide causal evidence supporting common perceptions about the effects of unionization on inequality that were previously difficult to obtain as credibly.

References

Abdulkadĭroğlu, Atĭla, Joshua D Angrist, Yusuke Narita, and Parag Pathak, "Breaking Ties: Regression Discontinuity Design Meets Market Design," *Econometrica*, 2022, *90* (1), 117–151.

Akcigit, Ufuk, Salomé Baslandze, and Francesca Lotti, "Connecting to power: political connections, innovation, and firm dynamics," *Econometrica*, 2023, *91* (2), 529–564.

Akey, Pat, "Valuing changes in political networks: Evidence from campaign contributions to close congressional elections," *The Review of Financial Studies*, 2015, *28* (11), 3188–3223.

Almond, Douglas, Joseph J Doyle Jr, Amanda E Kowalski, and Heidi Williams, "Estimating marginal returns to medical care: Evidence from at-risk newborns," *The Quarterly Journal of Economics*, 2010, *125* (2), 591–634.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2009.

Angrist, Joshua, Peter Hull, Parag A Pathak, and Christopher Walters, "Credible school value-added with undersubscribed school lotteries," *Review of Economics and Statistics*, 2024, *106* (1), 1–19.

Ater, Itai, Yael Elster, and Eran B Hoffmann, "Real-estate investors, house prices and rents: Evidence from capital-gains tax changes," *Maurice Falk Institute for Economic Research in Israel. Discussion paper series*, 2021, (1), 1–33.

Auerbach, Eric, Yong Cai, and Ahnaf Rafi, "Regression Discontinuity Design with Spillovers," *arXiv preprint arXiv:2404.06471*, 2024.

Autor, David, David Dorn, and Gordon Hanson, "When work disappears: Manufacturing decline and the falling marriage market value of young men," *American Economic Review: Insights*, 2019, *1* (2), 161–178.

Autor, David H, Lawrence F Katz, and Melissa S Kearney, "Trends in US wage inequality: Revising the revisionists," *The Review of economics and statistics*, 2008, *90* (2), 300–323.

Azoulay, Pierre, Joshua S Graff Zivin, Danielle Li, and Bhaven N Sampat, "Public R&D investments and private-sector patenting: evidence from NIH funding rules," *The Review of economic studies*, 2019, *86* (1), 117–152.

Bahar, Dany, Prithwiraj Choudhury, Do Yoon Kim, and Wesley W Koo, "Innovation on wings: Nonstop flights and firm innovation in the global context," *Management Science*, 2023, *69* (10), 6202–6223.

Baron, E Jason, "School spending and student outcomes: Evidence from revenue limit elections in Wisconsin," *American Economic Journal: Economic Policy*, 2022, *14* (1), 1–39.

Baskaran, Thushyanthan and Zohal Hessami, "Does the election of a female leader clear the way for more women in politics?," *American Economic Journal: Economic Policy*, 2018, *10* (3), 95–121.

_ and _ , "Women in political bodies as policymakers," *Review of Economics and Statistics*, 2023, pp. 1–46.

_ , Sonia Bhalotra, Brian Min, and Yogesh Uppal, "Women legislators and economic performance," *Journal of Economic Growth*, 2024, *29* (2), 151–214.

_ , **Zohal Hessami, and Sebastian Schirner**, "Young versus old politicians and public spending priorities," *Journal of Economic Behavior & Organization*, 2024, *225*, 88–106.

Beach, Brian and Daniel B Jones, "Gridlock: Ethnic diversity in government and the provision of public goods," *American Economic Journal: Economic Policy*, 2017, *9* (1), 112–136.

Ben-Porath, Netanel and Matan Kolerman-Shemer, "Political Priorities In A Competitive Landscape: Female Leaders and the Prevalence of Abortions," 2024.

Besley, Timothy and Anne Case, "Political institutions and policy choices: evidence from the United States," *Journal of Economic Literature*, 2003, *41* (1), 7–73.

Bhalotra, Sonia and Irma Clots-Figueras, "Health and the political agency of women," *American Economic Journal: Economic Policy*, 2014, *6* (2), 164–197.

_ , _ , **and Lakshmi Iyer**, "Pathbreakers? Women's electoral success and future political participation," *The Economic Journal*, 2018, *128* (613), 1844–1878.

_ , _ , **and** _ , "Religion and abortion: The role of politician identity," *Journal of Development Economics*, 2021, *153*, 102746.

_ , _ , **Guilhem Cassan, and Lakshmi Iyer**, "Religion, politician identity and development outcomes: Evidence from India," *Journal of Economic Behavior & Organization*, 2014, *104*, 4–17.

Biasi, Barbara, "The labor market for teachers under different pay schemes," *American Economic Journal: Economic Policy*, 2021, *13* (3), 63–102.

_ **and Heather Sarsons**, "Flexible wages, bargaining, and the gender gap," *The Quarterly Journal of Economics*, 2022, *137* (1), 215–266.

_ , **Julien M Lafortune, and David Schönholzer**, "What Works and for Whom? Effectiveness and Efficiency of School Capital Investments across the US," Technical Report, National Bureau of Economic Research 2024.

Black, Sandra E, "Do better schools matter? Parental valuation of elementary education," *The Quarterly Journal of Economics*, 1999, *114* (2), 577–599.

Borusyak, Kirill and Peter Hull, "Nonrandom exposure to exogenous shocks," *Econometrica*, 2023, *91* (6), 2155–2185.

_ , _ , **and Xavier Jaravel**, "Quasi-experimental shift-share research designs," *The Review of Economic Studies*, 2022, *89* (1), 181–213.

_ , _ , **and** _ , "A practical guide to shift-share instruments," Technical Report, National Bureau of Economic Research 2024.

Callaway, Brantly and William J Collins, "Unions, workers, and wages at the peak of the American labor movement," *Explorations in Economic History*, 2018, *68*, 95–118.

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik, "Robust nonparametric confidence intervals for regression-discontinuity designs," *Econometrica*, 2014, *82* (6), 2295–2326.

_ , _ , **Max H Farrell, and Rocio Titiunik**, "Regression discontinuity designs using covariates," *Review of Economics and Statistics*, 2019, *101* (3), 442–451.

Campante, Filipe and David Yanagizawa-Drott, "Long-range growth: economic development in the global network of air links," *The Quarterly Journal of Economics*, 2018, *133* (3), 1395–1458.

Card, David, "The effect of unions on the structure of wages: A longitudinal analysis," *Econometrica*, 1996, pp. 957–979.

_ , "The effect of unions on wage inequality in the US labor market," *ILR Review*, 2001, *54* (2), 296–315.

Carozzi, Felipe, Davide Cipullo, and Luca Repetto, "Political fragmentation and government stability: Evidence from local governments in Spain," *American Economic Journal: Applied Economics*, 2022, *14* (2), 23–50.

Cattaneo, Matias D and Rocio Titiunik, "Regression discontinuity designs," *Annual Review of Economics*, 2022, *14*, 821–851.

_ , **Brigham R Frandsen, and Rocio Titiunik**, "Randomization inference in the regression discontinuity design: An application to party advantages in the US Senate," *Journal of Causal Inference*, 2015, *3* (1), 1–24.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein, "The value of school facility investments: Evidence from a dynamic regression discontinuity design," *The Quarterly Journal of Economics*, 2010, *125* (1), 215–261.

Census, US, *1987 Economic Censuses: Report series. Release 1E. Technical documentation. CD-ROM volume 1E* number pt. 1. In '1987 Economic Censuses: Report Series. Release 1E. Technical Documentation. CD-ROM Volume 1E.', U.S. Department of Commerce, Economics and Statistics Administration, Bureau of the Census, 1993.

Chaisemartin, Clément De and Xavier d'Haultfoeuille, "Two-way fixed effects estimators with heterogeneous treatment effects," *American Economic Review*, 2020, *110* (9), 2964–2996.

Clark, Damon, "The performance and competitive effects of school autonomy," *Journal of political Economy*, 2009, *117* (4), 745–783.

Clots-Figueras, Irma, "Women in politics: Evidence from the Indian States," *Journal of public Economics*, 2011, *95* (7-8), 664–690.

_ , "Are female leaders good for education? Evidence from India," *American Economic Journal: Applied Economics*, 2012, *4* (1), 212–44.

Collins, William J and Gregory T Niemesh, "Unions and the Great Compression of wage inequality in the US at mid-century: evidence from local labour markets," *The Economic History Review*, 2019, *72* (2), 691–715.

Cooper, Michael J, Huseyin Gulen, and Alexei V Ovtchinnikov, "Corporate political contributions and stock returns," *The Journal of Finance*, 2010, *65* (2), 687–724.

Dahl, Gordon B, Katrine V Løken, and Magne Mogstad, "Peer effects in program participation," *American Economic Review*, 2014, *104* (7), 2049–2074.

Darolia, Rajeev, "Integrity versus access? The effect of federal financial aid availability on postsecondary enrollment," *Journal of Public Economics*, 2013, *106*, 101–114.

de Benedictis-Kessner, Justin and Christopher Warshaw, "Politics in forgotten governments: the partisan composition of county legislatures and county fiscal policies," *The Journal of Politics*, 2020, *82* (2), 460–475.

_ , Daniel Jones, and Christopher Warshaw, "How partisanship in cities influences housing policy," *American Journal of Political Science*, 2024.

Dechezleprêtre, Antoine, Elias Einiö, Ralf Martin, Kieu-Trang Nguyen, and John Van Reenen, "Do Tax Incentives Increase Firm Innovation? An RD Design for R&D, Patents, and Spillovers," *American Economic Journal: Economic Policy*, 2023, *15* (4), 486–521.

Dell, Melissa, "The persistent effects of Peru's mining mita," *Econometrica*, 2010, *78* (6), 1863–1903.

_ and Pablo Querubin, "Nation building through foreign intervention: Evidence from discontinuities in military strategies," *The Quarterly Journal of Economics*, 2018, *133* (2), 701–764.

DiNardo, John and David S Lee, "Economic impacts of new unionization on private sector employers: 1984–2001," *The Quarterly Journal of Economics*, 2004, *119* (4), 1383–1441.

Dube, Arindrajit, Laura Giuliano, and Jonathan Leonard, "Fairness and frictions: The impact of unequal raises on quit behavior," *American Economic Review*, 2019, *109* (2), 620–663.

Eckert, Fabian, Teresa C. Fort, Peter K. Schott, and Natalie J. Yang, "Imputing Missing Values in the US Census Bureau's County Business Patterns," Technical Report, National Bureau of Economic Research 2021.

Farber, Henry S, Daniel Herbst, Ilyana Kuziemko, and Suresh Naidu, "Unions and inequality over the twentieth century: New evidence from survey data," *The Quarterly Journal of Economics*, 2021, *136* (3), 1325–1385.

Feigenbaum, James, Alexander Hertel-Fernandez, and Vanessa Williamson, "From the Bargaining Table to the Ballot Box: Downstream Effects of Right-to-Work Laws," Technical Report, Working Paper 2019.

Ferreira, Fernando and Joseph Gyourko, "Does gender matter for political leadership? The case of US mayors," *Journal of Public Economics*, 2014, *112*, 24–39.

Folke, Olle, "Shades of brown and green: party effects in proportional election systems," *Journal of the European Economic Association*, 2014, *12* (5), 1361–1395.

Fortin, Nicole, Thomas Lemieux, and Neil Lloyd, "Labor market institutions and the distribution of wages: The role of spillover effects," *Journal of Labor Economics*, 2021, *39* (S2), S369–S412.

_ , _ , and _ , "Right-to-Work Laws, Unionization, and Wage Setting," Technical Report, National Bureau of Economic Research 2022.

Frandsen, Brigham R., "Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete," in "Regression discontinuity designs," Emerald Publishing Limited, 2017.

_ , "The surprising impacts of unionization: Evidence from matched employer-employee data," *Journal of Labor Economics*, 2021, *39* (4), 861–894.

Fu, Chao and Jesse Gregory, "Estimation of an equilibrium model with externalities: Post-disaster neighborhood rebuilding," *Econometrica*, 2019, *87* (2), 387–421.

Gelman, Andrew and Guido Imbens, "Why high-order polynomials should not be used in regression discontinuity designs," *Journal of Business & Economic Statistics*, 2019, *37* (3), 447–456.

Geys, Benny, Zuzana Murdoch, and Rune J Sørensen, "Public Employees as Elected Politicians: Assessing Direct and Indirect Substantive Effects of Passive Representation," *The Journal of Politics*, 2024, *86* (1), 170–182.

Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen J Redding, "Trade and inequality: From theory to estimation," *The Review of Economic Studies*, 2016, *84* (1), 357–405.

Holmes, Thomas J, "The effect of state policies on the location of manufacturing: Evidence from state borders," *Journal of political Economy*, 1998, *106* (4), 667–705.

Hsu, YC and S Shen, "Dynamic regression discontinuity under treatment effect heterogeneity," *Work. Pap., Univ. Calif., Davis*, 2021.

Hyytinen, Ari, Jaakko Meriläinen, Tuukka Saarimaa, Otto Toivanen, and Janne Tukiainen, "Public employees as politicians: Evidence from close elections," *American Political Science Review*, 2018, *112* (1), 68–81.

Imbens, Guido and Karthik Kalyanaraman, "Optimal bandwidth choice for the regression discontinuity estimator," *The Review of economic studies*, 2012, *79* (3), 933–959.

IPUMS USA, "1970 County Group Definitions: Composition File," 2024. Geographic Tools Documentation.

Isen, Adam, "Do local government fiscal spillovers exist? Evidence from counties, municipalities, and school districts," *Journal of Public Economics*, 2014, *110*, 57–73.

Jackson, C Kirabo and Claire L Mackevicius, "What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States," *American Economic Journal: Applied Economics*, 2024, *16* (1), 412–446.

Johnson, Matthew S, "Regulation by shaming: Deterrence effects of publicizing violations of workplace safety and health laws," *American economic review*, 2020, *110* (6), 1866–1904.

Jones, Maria, Florence Kondylis, John Loeser, and Jeremy Magruder, "Factor market failures and the adoption of irrigation in rwanda," *American Economic Review*, 2022, *112* (7), 2316–2352.

Keele, Luke and Rocío Titiunik, "Geographic natural experiments with interference: The effect of all-mail voting on turnout in Colorado," *CESifo Economic Studies*, 2018, *64* (2), 127–149.

Knepper, Matthew, "From the fringe to the fore: Labor unions and employee compensation," *Review of Economics and Statistics*, 2020, *102* (1), 98–112.

Kolerman-Shemer, Matan, "Political Impacts of Trade Unions," 2023.

Kott, Assaf, "Income Shocks, School Choice, and Long-Term Outcomes: Lessons from Child Allowances in Israel," 2022.

Kuliomina, Jekaterina, "Do personal characteristics of councilors affect municipal budget allocation?," *European Journal of Political Economy*, 2021, *70*, 102034.

Lang, Corey, "Assessing the efficiency of local open space provision," *Journal of Public Economics*, 2018, *158*, 12–24.

Lee, David S, "Randomized experiments from non-random selection in US House elections," *Journal of Econometrics*, 2008, *142* (2), 675–697.

_ **and Alexandre Mas**, "Long-run impacts of unions on firms: New evidence from financial markets, 1961–1999," *The Quarterly Journal of Economics*, 2012, *127* (1), 333–378.

Lewis, H. Gregg, *Unionism and Relative Wages in the United States*, Chicago: University of Chicago Press, 1963.

Makridis, Christos Andreas, "Do right-to-work laws work? evidence on individuals' well-being and economic sentiment," *The Journal of Law and Economics*, 2019, *62* (4), 713–745.

Manski, Charles F, "Identification of endogenous social effects: The reflection problem," *The review of economic studies*, 1993, *60* (3), 531–542.

Martorell, Paco, Kevin Stange, and Isaac McFarlin Jr, "Investing in schools: capital spending, facility conditions, and student achievement," *Journal of Public Economics*, 2016, *140*, 13–29.

Meriläinen, Jaakko, "Political selection and economic policy," *The Economic Journal*, 2022, *132* (648), 3020–3046.

Mora-García, Claudio A and Tomás Rau, "Peer Effects in the Adoption of a Youth Employment Subsidy," *Review of Economics and Statistics*, 2023, *105* (3), 614–625.

Narita, Yusuke and Kohei Yata, "Algorithm is experiment: Machine learning, market design, and policy eligibility rules," *arXiv preprint arXiv:2104.12909*, 2023.

National Labor Relations Board, "Basic Guide to the National Labor Relations Act: General Principles of Law Under the Statute and Procedures of the National Labor Relations Board," Government Document, National Labor Relations Board, Washington, DC 1997. Originally issued in 1962, revised edition.

Nellis, Gareth and Niloufer Siddiqui, "Secular party rule and religious violence in Pakistan," *American Political Science Review*, 2018, *112* (1), 49–67.

_ **, Michael Weaver, Steven C Rosenzweig et al.**, "Do parties matter for ethnic violence? Evidence from India," *Quarterly Journal of Political Science*, 2016, *11* (3), 249–277.

Obama, Barack, "President Obama's Open Letter to America's Hardworking Men and Women," https://obamawhitehouse.archives.gov/blog/2016/09/04/president-obama-letter-americas-hardworking-men-and-women September 2016. At 8:00 PM ET.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, "Distributional national accounts: methods and estimates for the United States," *The Quarterly Journal of Economics*, 2018, *133* (2), 553–609.

Priyanka, Sadia, "Do female politicians matter for female labor market outcomes? Evidence from state legislative elections in India," *Labour Economics*, 2020, *64*, 101822.

Saez, Emmanuel, Benjamin Schoefer, and David Seim, "Payroll taxes, firm behavior, and rent sharing: Evidence from a young workers' tax cut in Sweden," *American Economic Review*, 2019, *109* (5), 1717–1763.

Santoleri, Pietro, Andrea Mina, Alberto Di Minin, and Irene Martelli, "The causal effects of R&D grants: Evidence from a regression discontinuity," *Review of Economics and Statistics*, 2022, pp. 1–42.

Small, Dylan S, Zhiqiang Tan, Roland R Ramsahai, Scott A Lorch, and M Alan Brookhart, "Instrumental variable estimation with a stochastic monotonicity assumption," 2017.

Sojourner, Aaron J, Brigham R Frandsen, Robert J Town, David C Grabowski, and Min M Chen, "Impacts of unionization on quality and productivity: Regression discontinuity evidence from nursing homes," *ILR Review*, 2015, *68* (4), 771–806.

Sørensen, Rune J, "Educated politicians and government efficiency: Evidence from Norwegian local government," *Journal of Economic Behavior & Organization*, 2023, *210*, 163–179.

Tan, Brandon Joel, "The consequences of letter grades for labor market outcomes and student behavior," *Journal of Labor Economics*, 2023, *41* (3), 565–588.

Torrione, Elena Dal, Tiziano Arduini, and Laura Forastiere, "Regression Discontinuity Designs Under Interference," *arXiv preprint arXiv:2410.02727*, 2024.

Vaghul, Kavya and Ben Zipperer, "Historical state and sub-state minimum wage data," 2016.

Valentim, Vicente and Elias Dinas, "Does party-system fragmentation affect the quality of democracy?," *British Journal of Political Science*, 2024, *54* (1), 152–178.

Wang, Sean and Samuel Young, "Unionization, Employer Opposition, and Establishment Closure," *Essays on Employment and Human Capital, PhD diss. MIT*, 2022.

Yitzhaki, Shlomo, "On using linear regressions in welfare economics," *Journal of Business & Economic Statistics*, 1996, *14* (4), 478–486.

Zigler, Corwin M and Georgia Papadogeorgou, "Bipartite causal inference with interference," *Statistical science: a review journal of the Institute of Mathematical Statistics*, 2021, *36* (1), 109.

Appendix

A Proof of Proposition 2

We start with the lower-level solution. Assuming that the population coefficients on r_j and r_j^+ in the first-stage and reduced-form of (9) have some limits as $h \to 0$, including them in the specification does not affect the set of estimands, as both of these variables have no variation in the limit. Thus, the estimand limit is

$$\lim_{h \downarrow 0} \beta_h^\ell = \frac{\lim_{r \downarrow 0} \mathbb{E}\left[s_j Y_{\mathbf{i}(j)} \mid r_j = r\right] - \lim_{r \uparrow 0} \mathbb{E}\left[s_j Y_{\mathbf{i}(j)} \mid r_j = r\right]}{\lim_{r \downarrow 0} \mathbb{E}\left[s_j X_{\mathbf{i}(j)} \mid r_j = r\right] - \lim_{r \uparrow 0} \mathbb{E}\left[s_j X_{\mathbf{i}(j)} \mid r_j = r\right]}.$$

We can expand $\mathbb{E}\left[Y_{\mathbf{i}(j)} \mid r_j = r\right] = \mathbb{E}\left[Y_{\mathbf{i}(j)}(X_{\mathbf{i}(j)}(1, \boldsymbol{z}_{\mathbf{i}(j)-j})) \mid r_j = r\right]$ for $r \geq 0$, with analogous expressions for other terms. Thus, by Assumption 3a,

$$\beta_h^\ell = \frac{\mathbb{E}\left[s_j Y_{\mathbf{i}(j)}(X_{\mathbf{i}(j)}(1, \boldsymbol{z}_{\mathbf{i}(j)-j})) \mid r_j = 0\right] - \mathbb{E}\left[s_j Y_{\mathbf{i}(j)}(X_{\mathbf{i}(j)}(0, \boldsymbol{z}_{\mathbf{i}(j)-j})) \mid r_j = 0\right]}{\mathbb{E}\left[s_j X_{\mathbf{i}(j)}(1, \boldsymbol{z}_{\mathbf{i}(j)-j}) \mid r_j = 0\right] - \mathbb{E}\left[s_j X_{\mathbf{i}(j)}(0, \boldsymbol{z}_{\mathbf{i}(j)-j}) \mid r_j = 0\right]},$$

where conditioning on $r_j = 0$ is well-defined by Assumption 4. By Assumption 5, the denominator is positive. Moreover, by standard arguments (e.g., Yitzhaki (1996)), it is a convexly-weighted average of $\partial Y_i(x)/\partial x$ (or the discrete analog thereof when X_i is discrete) across both units and values of treatment.

By Proposition 1, the upper-level solution only differs by residualizing Y_i and X_i on Q_i. By Assumption 3b, we have $Q_i \approx (s_j, 0, 0)$ conditionally on $-h \leq r_j \leq h$ as $h \to 0$. Invoking Assumption 3b again and assuming again that the population coefficients on Q_i in the first-stage and reduced-form of (6) converge, this implies that residualizing on Q_i makes no difference in (8), provided an intercept is included in that subunit-level specification. This concludes the proof.

B Data Appendix

Our analysis is at the level of state-industry cells for each decennial census year from 1960–2010 (i.e., each year that divides by ten). We begin by discussing the data construction steps for our three data sources separately: decennial censuses, CPS, the unionization elections data. We then turn to how we merge them into a consistent cell-level panel; this involves creating a harmonized industry classification and computing cell-level variables, such as inequality measures.

Decennial census data. Census public-use micro-data were downloaded from the IPUMS USA website. For 1960, 1980, 1990, and 2000, we used the 5% "state" samples. For 1970, we combined the 2% "state" and 2% "metro" samples. To represent 2010, we combine the American Community Survey (ACS) samples for 2009–2013. All of these samples are nationally representative for the US population.

We assign each individual to their state of residence. While state of employment could be more relevant for our analysis, it is not available in some of the samples. The state of residence is observed in most samples; the only exception is the 1970 metro samples which do not report the state for the 17% of individuals living in metropolitan areas and other county groups that cross state boundaries. We duplicate observations for those individuals and adjust their projection weights to equal the original weight multiplied by each state's

proportion of the metropolitan area or county group's total population.[41] We discuss the assignment of workers to industries below.

For each worker, we observe their pre-tax annual labor income (variable `INCWAGE`). Top-coded income values are multiplied by 1.5 (Autor et al., 2008). We then compute the hourly wage as the income dividing by the weeks of work and the average number of hours per week.

We measure workers' education and potential work experience using the most detailed education question (`EDUCD`). The groupings defined by this question vary across samples, sometimes closer aligned with the years of schooling and other times with the highest degree obtained. In all cases, it is sufficiently detailed to determine whether the individual has exactly a high school degree, exactly a college degree, or other education levels above and below. It is also sufficient to measure years of schooling, at least approximately, which we use define potential work experience as the age minus 6 and minus the years of schooling.

We classify a worker's occupation as managerial if the occupation variable (`OCC1990`) belongs to the groups "Executive, Administrative, and Managerial" and "Management Related" .

We impose several sample restrictions. First, we focus on 18–65 year-old adults with 0–48 years of potential experience. Second, we only keep workers who have been employed for all 52 weeks during the year preceding the survey and worked at least 20 hours a week on average.[42] Individuals who worked fewer weeks likely changed jobs within the year, making it more difficult to correctly assign them to an industry. Third, we require that workers earn at least half their state's minimum hourly wage (obtained from Vaghul and Zipperer (2016)). Fourth, we consider only workers in the private sector who are not self-employed; those are the workers who can unionize through NLRB unionization elections.

CPS data. To capture variables not available in the census—unionization rates and fringe benefits—we leverage CPS samples, downloaded from IPUMS CPS.[43] To measure unionization rates, we use CPS MORG (Merged Outgoing Rotation Group) samples from 1983–2015, along with the May supplement to the CPS in 1977–1980. The surveys ask workers whether they are union members (`UNION` variable). For fringe benefits, we use CPS ASEC (Annual Social and Economic Supplements) from 1980–2014, along with the May supplement in 1979. The workers are asked if they are covered by an employer-sponsored pension plan (`PENSION` variable) and whether they enrolled in their employer's group health plan during the previous year (`INCLUGH` variable). We use the same four individual-level sample restrictions as in the decennial census. Imputed answers were dropped.

To increase sample size when measuring changes in the rates of unionization, pension, and insurance coverage between adjacent census years, we pool data from multiple years of the CPS. Where possible, we compare the five years after the end of the decade to five years before the beginning of the decade. For instance, the differenced outcome for the decade of the 2000s is calculated as the change between 1986–90 and 2000–04. We deviate from this rule in the 1980s because of data availability; Appendix Table A9 provides the details.

Unionization elections. We use the NLRB (National Labor Relations Board) unionization election data maintained by Hank Farber and John-Paul Ferguson. We keep observations from 1961—the earliest available

[41] We used IPUMS USA (2024) to identify the county composition of each metropolitan area and county group, and obtained the 1970 county populations from IPUMS NHGIS.

[42] Since 1980, the variable is based on the `UHRSWORK` variable that measures "usual hours worked per week [in the last year]." Unfortunately, this variable is not available in the 1970 and the 1960 samples, so for those years we use the `HRSWORK2` variable that measures "hours worked last week" instead.

[43] With the exception of the May supplement for 1979, which was downloaded from NBER.

year—until the beginning of 2009, after which industry variables are no longer available. The original dataset includes 250,063 elections.

Each election is attributed to a geographic location (directly including the state), an industry, and a decade. We generally define decades as starting in the year that ends with 0 and finishing in the year which ends with 9, e.g. 1980–1989. The earliest and latest decades are slightly shorter because, as just mentioned, the data start in 1961 and end early in 2009. We discuss industry classifications below.

We observe the number of eligible voters in the bargaining unit, the number of votes cast for and against the union, as well as whether the union has been formed. In most cases, the union is formed when the vote share (out of the total votes cast) strictly exceeds 50%, but in a small fraction of cases this is not the case, perhaps because of the reversal during the appeal process. We further observe whether the election is a union decertification election. In the main analysis, we do not distinguish between certification and decertification elections, counting workers in re-certified unions as newly unionized; we make other choices in a robustness check, as well as in Section 5.5.

While we use the full set of union elections to construct the treatment variable,[44] our instrument is constructed only from the elections that satisfy several criteria. Of course, it only includes close elections, defined by the vote share in favor of the union falling within a bandwidth around the cutoff of 50%. We make two additional restrictions. First, following common practice in the unionization RD literature, elections with fewer than 20 votes were excluded (DiNardo and Lee, 2004; Lee and Mas, 2012; Frandsen, 2021). Second, we drop elections that ended in a tie. Although by the National Labor Relations Act a tie is equivalent to a union loss, including those elections in the sample may lead to bias. Elections with an even number of votes will mechanically have a much higher chance of resulting in a close union loss than in a close union win. For example, consider an election with 20 votes given bandwidth of 10%. There are there outcomes resulting in a close loss (8-12, 9-11, and 10-1) but only two outcomes resulting in a close win (12-8 and 11-9). Excluding the tie, we prevent a negative mechanical correlation between the election result and workplace size.[45]

Third, following Frandsen (2017, 2021), the instrument construction also excludes elections that ended with a one-vote margin in either direction. Frandsen (2017, 2021) documented the imbalance of covariates and unequal frequency of union wins and losses in such elections. Figure A3 confirms this in our dataset, showing approximately 18,000 "missing" workers in elections that ended with exactly a one-vote margin in favor of the unions. A presumed explanation is that published unionization result may in some cases report vote counts post appeals and other legal process, and that the tendency to appeal and the acceptance rate of those appeals is heterogeneous between close wins and close losses.

Table A2 presents summary statistics for the full set of unionization elections, as well as the restricted set for the instrument construction. The average vote share for the unions and the share of union wins are close to 50% in both samples. The average number of votes in the treatment (instrument) sample is 54.9 (98.3), and the share of manufacturing establishments is 43% (52%). In both samples, over half of the elections occurred in the 1960s and 1970s, which constitute only 40% of the study period, reflecting the decline in unionization since 1980.

Industry classification. To construct both labor market outcomes and new unionization rates by state-industry cells, we had to create an industry classification harmonized both across data sources and over time.

[44] Except a small fraction that we cannot attribute to an industry. This explains the slightly smaller total number of elections of 232,267 in Table A2. We include the entire set of 250,063 elections in the counterfactual analysis of Section 5.5.

[45] In support of this argument, Wang and Young (2022) provides empirical evidence that tied elections are very different from elections where the union lost or won by a slightly larger margin.

The classifications in all our data sources are based, in one way or another, on SIC codes before around 2000 and on NAICS codes thereafter. Thus, we also created two harmonized classifications for the two subperiods, resulting in 70 industries based on 2-digit SIC87 codes and 73 industries based on 3-digit NAICS97 codes covering the entire private sector of the economy. We use our SIC-based classification for the 1960s–1990s and the NAICS-based classification for the 2000s.

For union elections filed before the end of 1998 (which covers most elections that took place before the end of 1999), the NLRB data report the 3-digit SIC code (with some changes in the classification over time discussed below). The decennial censuses up until 1990 and the CPS up until 2002 use a different classification—variable `IND1990`—that was harmonized over time by IPUMS.[46] This classification is loosely based on on 3-digit SIC87 codes. To minimize measurement error, we convert `IND1990` industries to 2-digit SIC87 codes by leveraging crosswalks created by David Dorn (Autor et al., 2019). Specifically, he published many-to-one crosswalks from `IND1990` to a slightly aggregated classification developed by him called `IND1990ddx`, as well as from 4-digit SIC87 codes to `IND1990ddx`. While the resulting mapping from `IND1990` to the 3-digit SIC87 codes that we observe in the NLRB data would be many-to-many, the mapping to 2-digit SIC87 codes is mostly many-to-one. In the rare cases where this is not the case, we merge those 2-digit SIC87 codes together, resulting in the 70 industry codes in our classification.

We make adjustments for the changes in the SIC classifications over time in the NLRB data. While SIC87 codes are reported since 1987, the SIC72 classification is used between 1972 and 1986. In those years, we only keep the elections in industries for which over 95% of national employment mapped to one of our (significantly more aggregated) codes.[47] Elections prior to 1972 use even earlier SIC schemes. We drop all elections that did not align with a single code in our classification using the crosswalk used by Eckert et al. (2021)..

The NAICS-based concordance is more straightforward. For union elections filed since 1999, the NLRB data report the 3-digit NAICS code, also with some changes in the classification over time. The decennial censuses of 2000 and 2010 and the CPS starting in 2003 report the variable `INDNAICS` which is closely based on NAICS97 codes and harmonized over time. Specifically, its first three digits coincide with 3-digit NAICS97, except for "not specified" industries which we drop and a few other industries which we merge together.[48] We also merge a few codes where the NAICS codes changed over time in the decennial censuses and CPS in a way that is not many-to-one from the newer codes to NAICS97, ultimately resulting in our 73-industry classification. In total, across the entire period of our study, we were able to match 92.9% of unionization elections and 99% of census records to our industry codes.

The transition from SIC to NAICS industries creates additional complications when measuring outcomes in 10-year changes. For decennial census variables, this affects the change from 1990 to 2000. The 2000 census, however, includes the `IND1990` variable converted from `INDNAICS` using a crosswalk designed by IPUMS. We therefore compute the 1990 to 2000 change using our SIC-based classification. While the conversion could introduce noise in differenced outcomes, we find it reassuring that employment changes do not exhibit substantially higher volatility in the 1990s than in the decade before or after.[49]

[46] Note that the harmonization process conducted by IPUMS follows a majority rule, assigning each original industry code to the harmonized industry that the majority of observations would have been assigned to. This process may lead to some inconsistencies between the census and the unionization data in earlier decades.

[47] This calculation relies on the economic censuses of 1987 (Census, 1993), which report total employment data according to both SIC72 and SIC87.

[48] We dropped codes 3MS "Not specified industries" and 4MS "Not specified retail trade." An example of industries that we merged is construction, which has only one `INDNAICS` code (code 22) but multiple 3-digit NAICS codes.

[49] A small issue also arises from the fact that about one-third of the unionization elections that happened in 1999 are recorded with NAICS codes. We exclude these elections from the RDA variables construction.

Industry transitions in the CPS, which happened between 2002 and 2003, require more care. This affects differenced outcomes not only in the 1990s (measured as 2000–2004 minus 1986–1990) but also in the 2000s (measured as 2010–2014 minus 1996–2000; see Table A9). The conversion from `INDNAICS` codes to `IND1990` for 2003–04 was done by IPUMS, and we use it to compute the changes over the 1990s in our SIC-based classification. In the other direction, we apply the IPUMS crosswalk ourselves, converting `IND1990` to `INDNAICS` for 1996–2000 to compute the changes over the 2000s using our NAICS-based classification.

State-by-industry panel and variable construction. Once all of our datasets have harmonized state and industry information and are assigned to a decade (in the NLRB data) or a census year (in the decennial censuses and the CPS), we compute differenced outcomes in the panel of state-level cells by decade. Since our industry classification changes in 2000, we compute all outcomes for the year 2000 using both classifications.

The computation of the treatment, instrument, and the RDA controls is described in detail in the main text. We drop a small number of cell-decades in which there were more workers involved in unionization attempts than the total number of workers at the beginning of the decade according to the census, as those cells probably contain miscoding of the industry variable or inconsistency between the census and NLRB coding.

Most outcomes are straightforward. The union density, pension coverage, and insurance coverage rates are fractions of the corresponding answers in the CPS, using projection weights. In the decennial censuses, the Gini index and the top-10% income share are calculated based on the total labor income. The log 90-10 wage ratio and the variance of log wages, along with average wages for different workforce segments, are computed based from log hourly wages. All measures are weighted by projection weights. Employment is measured as the sum of weights within each cell.

To estimate the college wage premium by cell and year, we broadly follow Autor et al. (2008) and Farber et al. (2021). We select the sample or workers with either exactly a high-school degree or at least a college degree (or more). We then estimate the following regression by OLS with projection weights:

$$\log Wage_o = \gamma_{s(o)i(o)t(o)} College_o + \beta'_{s(o)t(o)} X_o + \text{error}_o \tag{21}$$

where o stands for an individual observation that belongs to state $s(o)$, industry $i(o)$, and year $t(o)$, $College_o$ is the dummy of having at least a college degree, and X_o include the following set of demographics variables: a female dummy, a non-white dummy, a quartic term in experience, and interaction terms of the female dummy with both the non-white dummy and the quartic in experience term. Like Farber et al. (2021), we allow the demographic coefficients to vary by state and decade. However, we restrict them to be the same across industries within each state-decade because the sample size would be insufficient for a fully-flexible specification. The estimates of γ_{sit} are the estimated college wage premiums.

For the inequality variables (except the college wage premium), overall wages and employment variables, and all CPS variables, we replace a cell-decade observation as missing if it represents less than 1,000 workers after applying projection weights or is based on less than 10 raw observations, either at the beginning or at the end of the decade. For census variables based on a subgroup, we require to have 10 raw observations for that subgroup: e.g., we only compute the average wage of college-educated workers in a cell if there are at least 10 college-educated respondents in that cell. For the college wage premium, we require at least 10 observations for both high school graduates and workers with a college degree or more.

Appendix Tables

Table A1: Published Papers Suitable for the RDA Design

Citation	Journal	Equations	Tables
Panel A: Difference-in-Differences Specifications			
Besley and Case (2003)	Journal of Economic Literature		
Cooper et al. (2010)	The Journal of Finance		
Saez et al. (2019)	American Economic Review		
Akcigit et al. (2023)	Econometrica		
Panel B: Aggregated Outcomes, Upper-Level Specifications			
Clots-Figueras (2011)	Journal of Public Economics	3,4	7
Clots-Figueras (2012)	American Economic Journal: Applied Economics	1,2	2,3
Bhalotra and Clots-Figueras (2014)	American Economic Journal: Economy Policy	1,2	2
Bhalotra et al. (2014)	Journal of Economic Behavior & Organization	1,2	2,3
Folke (2014)	Journal of the European Economic Association	4,5	3
Akey (2015)	The Review of Financial Studies	–	7
Nellis et al. (2016)	Quarterly Journal of Political Science	2,3	3
Nellis and Siddiqui (2018)	American Political Science Review	2,3	2
Campante and Yanagizawa-Drott (2018)	Quarterly Journal of Economics	7	4
Hyytinen et al. (2018)	American Political Science Review	1,4	3
Azoulay et al. (2019)	The Review of Economic Studies	3.3, 3.4	8
Priyanka (2020)	Labour Economics	1,2	3
Bhalotra et al. (2021)	Journal of Development Economics	1,2	2
Meriläinen (2022)	The Economic Journal	2,4	3
Bahar et al. (2023)	Management Science	6,7	8
Sørensen (2023)	Journal of Economic Behavior & Organization	1,2	2
Baskaran et al. (2024b)	Journal of Economic Behavior & Organization	1,2	1
Geys et al. (2024)	The Journal of Politics	1,4	Fig. 2
Valentim and Dinas (2024)	British Journal of Political Science	–	3
Panel C: Aggregated Outcomes, Lower-Level Specifications			
Ferreira and Gyourko (2014)	Journal of Public Economics	4	11
Beach and Jones (2017)	American Economic Journal: Economy Policy	1	3
Dell and Querubin (2018)	Quarterly Journal of Economics	1	2,3
de Benedictis-Kessner and Warshaw (2020)	Journal of Politics	–	Fig. 7
Carozzi et al. (2022)	American Economic Journal: Applied Economics	2,3	2
Baskaran and Hessami (2023)	The Review of Economics and Statistics	2	3
Tan (2023)	Journal of Labor Economics	2	2–4
de Benedictis-Kessner et al. (2024)	American Journal of Political Science	–	Fig. 4
Panel D: Spillovers			
Clark (2009)	Journal of Political Economy	2	4
Dahl et al. (2014)	American Economic Review	3–5	1
Isen (2014)	Journal of Public Economics	5,6	3–5
Baskaran and Hessami (2018)	American Economic Journal: Economic Policy	2	9
Bhalotra et al. (2018)	The Economic Journal	1	3
Dube et al. (2019)	American Economic Review	9	3
Fu and Gregory (2019)	Econometrica	4	Fig. 4
Johnson (2020)	American Economic Review	2	2
Santoleri et al. (2022)	The Review of Economics and Statistics	1	7
Jones et al. (2022)	American Economic Review	3	7,8
Mora-Garćia and Rau (2023)	The Review of Economics and Statistics	3,4	1
Baskaran et al. (2024a)	Journal of Economic Growth	1	4
Dechezleprêtre et al. (2023)	American Economic Journal: Economic Policy	5–7	6
Panel E: RD Aggregation over Time			
Cellini et al. (2010)	Quarterly Journal of Economics	12	4
Darolia (2013)	Journal of Public Economics	4,5	4–8
Martorell et al. (2016)	Journal of Public Economics	3	3
Lang (2018)	Journal of Public Economics	11	3
Baron (2022)	American Economic Journal: Economy Policy	2	5
Kuliomina (2021)	European Journal of Political Economy	–	4
Biasi et al. (2024)	NBER Working Paper	4	3

Notes: This table lists published papers featuring research designs suitable for RDA. We have organized these papers into the five categories as in Sections 3.1–3.3 and 4.1–4.2, although some papers may belong to more than one category. The "Equations" column specifies the regression specification most suitable for the RDA analysis, while the "Tables" column references the tables and figures containing the main results from these specifications. There are no references to equations and tables in Panel A since those papers do not use RDA-type specifications in settings where that could be possible.

Table A2: Unionization Data Summary Statistics

Variable	Treatment sample			Instrument sample		
	All	Union wins	Union losses	All	Union wins	Union losses
Number of elections	232267	116467	115800	118817	53024	61244
Union win	0.500	0.995	0.001	0.457	0.996	0.001
Union vote share	0.538	0.780	0.295	0.511	0.729	0.318
Number of votes	54.926	46.003	63.900	98.359	89.113	109.930
Number of eligible voters	64.804	57.874	71.774	116.123	113.187	123.262
Years 1961–1969	0.226	0.261	0.191	0.226	0.263	0.193
Years 1970–1979	0.326	0.316	0.335	0.314	0.297	0.327
Years 1980–1989	0.199	0.173	0.226	0.195	0.172	0.217
Years 1990–1999	0.150	0.141	0.159	0.162	0.153	0.169
Years 2000–2009	0.099	0.109	0.089	0.103	0.115	0.094
Manufacturing industry	0.430	0.410	0.449	0.520	0.500	0.539

Notes: The table reports mean characteristics for union elections included in our sample. The set of elections used to construct the treatment variable ("treatment sample") contains all elections, while the set used for the instrument construction ("instrument sample") contains only elections with the vote share for the union of $50 \pm 10\%$, at least 20 votes, and vote margin of at least 2 votes. Union wins are defined as strictly more than 50% of votes in favor of the union.
Sources: NLRB data.

Table A3: Within-Between Variance Decomposition of Log Wages, Census Data

Year	Var(log hourly wage)			Industry
	Total	Within cells	Between cells	classification
1960	0.262	0.192	0.071	SIC
1970	0.283	0.222	0.061	SIC
1980	0.284	0.228	0.056	SIC
1990	0.361	0.293	0.068	SIC
2000	0.387	0.326	0.061	SIC
2000	0.387	0.321	0.066	NAICS
2010	0.437	0.344	0.093	NAICS

Notes: The table presents a variance decomposition of log wages, broadly following Helpman et al. (2016). For each census year, we decompose the overall variance in log wages among workers in our Census sample into within and between (state-industry) cell components. In 2000, we perform the decomposition twice, using two industry coding frameworks: SIC-based (used in our main analysis until 2000) and NAICS-based (used from 2000 onwards). Census projection weights are used.
Sources: Decennial Census; authors' calculations.

Table A4: Instrument Balance, 15% Bandwidth

	NewUnions	Z		
	(1)	(2)	(3)	(4)
Log avg hourly wage$_{t-1}$	1.056	0.188	-0.155	-0.013
	(0.517)	(0.127)	(0.070)	(0.032)
Log college premium$_{t-1}$	4.452	2.331	-0.002	0.046
	(0.405)	(0.169)	(0.091)	(0.045)
Log 90/10 ratio$_{t-1}$	3.895	1.297	0.052	-0.005
	(0.630)	(0.250)	(0.133)	(0.062)
Gini index$_{t-1}$	-58.022	-20.529	-1.533	0.015
	(6.424)	(1.915)	(1.034)	(0.456)
Top 10% share$_{t-1}$	40.646	13.759	1.036	0.058
	(6.474)	(1.796)	(0.983)	(0.433)
Var(log hourly wage)$_{t-1}$	-7.027	-2.406	0.204	-0.162
	(1.486)	(0.387)	(0.219)	(0.100)
Log employment$_{t-1}$	-0.158	-0.095	-0.002	0.000
	(0.051)	(0.019)	(0.010)	(0.005)
Share with college degree$_{t-1}$	-4.510	-1.497	0.153	0.090
	(0.609)	(0.158)	(0.089)	(0.050)
Share white$_{t-1}$	-0.527	-0.302	-0.456	-0.049
	(0.561)	(0.241)	(0.132)	(0.051)
Share male$_{t-1}$	-0.869	-0.556	-0.078	0.037
	(0.516)	(0.170)	(0.082)	(0.036)
# significant controls	9	9	3	0
Partial R^2	0.1353	0.1049	0.0017	0.0000
Partial F-statistic	292.7	226.4	6.3	1.0
Partial F-test, p-value	0.0000	0.0000	0.0000	0.4557
RDA controls	None	None	$\sum_{j \in \mathcal{C}_{sit}} s_j$	All
Observations	12,799	12,799	12,799	12,799

Notes: Same as Table 2, except the instrument Z is constructed from close elections defined by the bandwidth of 15%, i.e. the union vote share of 50 ± 15 percent.
Sources: Decennial Census, NLRB data, and authors' calculations.

Table A5: Placebo Tests: The Impact of New Unionization on Predetermined Outcomes

Panel A: Lagged Outcomes in Levels

	College premium	Log 90/10	Gini coefficient	Top 10% share	Var(Log Wage)
Z_{sit}	0.046	0.052	-0.019	-0.018	0.013
	(0.207)	(0.225)	(0.059)	(0.047)	(0.077)
	[-0.736,0.760]	[-0.482,1.137]	[-0.211,0.235]	[-0.170,0.179]	[-0.225,0.313]
Observations	10,716	12,896	12,896	12,896	12,896

Panel B: Lagged Outcomes in Differences

	ΔCollege premium	ΔLog 90/10	ΔGini coefficient	ΔTop 10% share	ΔVar(Log Wage)
Z_{sit}	-0.344	0.491	0.038	-0.008	0.098
	(0.335)	(0.237)	(0.060)	(0.056)	(0.082)
	[-1.547,0.852]	[-0.279,1.391]	[-0.045,0.391]	[-0.122,0.309]	[-0.08,0.460]
Observations	6,467	7,706	7,706	7,706	7,706

Panel C: Covariates

	log(Avg wage)	log(Empl.)	Share college	Share white	Share male
Z_{sit}	-0.016	-0.587	-0.033	-0.062	0.096
	(0.178)	(1.180)	(0.065)	(0.112)	(0.095)
	[-0.472,1.098]	[-7.612,4.531]	[-0.286,0.250]	[-0.330,0.586]	[-0.299,0.611]
Observations	12,896	12,896	12,896	12,896	12,896

Notes: Placebo tests based on OLS estimates corresponding to the reduced-form of our main upper-level IV specification as in Table 3, Panel A. See notes to Table 3 for details about data and variables construction. The outcomes are measured at the beginning of the decade in which the instrument is measured in Panels A and C and as a 10-year lagged difference in Panel B. *Sources:* Decennial Census and NLRB unionization data; authors' calculations.

Table A6: The Impact of New Unionization on 10-Year Changes in Inequality: Robustness Checks

	ΔLog college premium (1)	ΔLog 90/10 (2)	ΔGini coeff. (3)	ΔTop 10% Share (4)	ΔVar(Log Wage) (5)
Panel A: Baseline					
Share newly unionized	-0.314	-0.459	-0.176	-0.143	-0.255
	(0.285)	(0.225)	(0.061)	(0.055)	(0.084)
	[-0.921,0.416]	[-0.933,0.122]	[-0.325,-0.035]	[-0.279,-0.020]	[-0.463,-0.054]
Panel B: 15% Bandwidth					
Share newly unionized	-0.361	-0.281	-0.123	-0.101	-0.198
	(0.211)	(0.151)	(0.041)	(0.038)	(0.057)
	[-0.856,0.095]	[-0.613,0.085]	[-0.222,-0.030]	[-0.193,-0.018]	[-0.33,-0.062]
Panel C: Excluding Union Decertification Elections					
Share newly unionized	-0.332	-0.758	-0.193	-0.147	-0.334
	(0.322)	(0.268)	(0.070)	(0.063)	(0.099)
	[-1.062,0.498]	[-0.987,0.169]	[-0.341,-0.023]	[-0.286,-0.003]	[-0.480,-0.038]
Panel D: State, Year, Industry Fixed Effects					
Share newly unionized	-0.393	-0.466	-0.165	-0.136	-0.220
	(0.275)	(0.237)	(0.062)	(0.055)	(0.089)
	[-0.930,0.341]	[-0.993,0.103]	[-0.311,-0.017]	[-0.269,-0.008]	[-0.433,-0.009]
Panel E: No Fixed Effects					
Share newly unionized	-0.416	-0.486	-0.180	-0.155	-0.261
	(0.335)	(0.265)	(0.070)	(0.062)	(0.102)
	[-1.126,0.473]	[-1.089,0.152]	[-0.336,-0.018]	[-0.297,-0.013]	[-0.503,-0.016]
Panel F: Unweighted					
Share newly unionized	-0.230	-0.518	-0.239	-0.187	-0.339
	(0.334)	(0.249)	(0.081)	(0.077)	(0.114)
	[-1.738,0.503]	[-1.089,0.400]	[-0.451,0.007]	[-0.421,0.014]	[-0.665,-0.035]

Notes: Panel A reports upper-level IV estimates of the effects of new unionization rates on wage inequality, replicating Panel A of Table 3. The following columns deviate from Panel A in different ways, one at a time. Panel B redefines the instrument and RDA controls using a bandwidth of 15% instead of 10%. Panel C reports results excluding decertification elections from the unionization elections sample when constructing the treatment, instrument, and RDA controls. Panel D includes separate state, industry, and year fixed effects instead of state-year and industry-year fixed effects. Panel E reports estimates without any fixed effects. Panel F presents unweighted estimation results, instead of weighting state-industry-decade observations by beginning-of-decade employment.
Sources: Decennial Census and NLRB unionization data; authors' calculations.

Table A7: The Impact of New Unionization on 10-Year Employment Changes

	All (1)	HS (2)	Non-College (3)	College (4)	College+ (5)
Panel A: Upper-Level Estimator					
Share newly unionized	-0.336	0.195	-0.148	-1.782	-1.719
	(0.474)	(0.489)	(0.459)	(0.895)	(0.829)
	[-1.638,0.701]	[-1.184,1.204]	[-1.381,0.872]	[-4.161,0.214]	[-3.961,0.085]
Mean outcome	10.987	9.960	10.678	8.943	9.292
Mean treatment	2.76%	2.76%	2.76%	2.76%	2.76%
RDA controls	✓	✓	✓	✓	✓
Industry-decade and state-decade FE	✓	✓	✓	✓	✓
Observations	12,910	12,910	12,910	12,723	12,792
Panel B: Lower-Level Estimator					
Share newly unionized	0.025	0.417	0.154	-1.086	-1.003
	(0.395)	(0.429)	(0.386)	(0.813)	(0.753)
	[-1.029,0.807]	[-0.736,1.236]	[-0.853,0.938]	[-3.402,0.347]	[-3.220,0.277]
Mean outcome	10.618	9.687	10.432	8.162	8.519
Mean treatment	9.54	9.54	9.54	9.53	9.54
RDA controls	✓	✓	✓	✓	✓
Industry-decade and state-decade FE	✓	✓	✓	✓	✓
Observations	31,256	31,256	31,256	31,167	31,207

Notes: This table reports IV estimates from the same specifications as in Table 3 but for different outcomes. The outcome in column 1 is the 10-year change in the total log employment in the state-industry cell. The other columns consider changes in log employment by education group: those with exactly a high-school degree (column 2), any level of education below a college degree (column 3), exactly a college degree (column 4), and college degree or more (column 5).
Sources: Decennial Census and NLRB unionization data; authors' calculations.

Table A8: The Impact of New Unionization on 10-Year Wage Changes: Robustness Checks

	All Workers (1)	College (2)	HS (3)	P90 (4)	P50 (5)	P10 (6)	Mgmt (7)	Non-Mgmt (8)
Panel A: Baseline								
Share newly unionized	-0.348	-0.449	-0.105	-0.282	-0.129	0.176	-0.917	-0.181
	(0.150)	(0.349)	(0.142)	(0.180)	(0.145)	(0.166)	(0.399)	(0.134)
	[-0.762,-0.032]	[-1.221,0.468]	[-0.442,0.227]	[-0.735,0.138]	[-0.533,0.188]	[-0.283,0.497]	[-1.854,0.058]	[-0.561,0.091]
Panel B: 15% Bandwidth								
Share newly unionized	-0.208	-0.340	0.079	-0.144	-0.052	0.136	-0.467	-0.077
	(0.102)	(0.252)	(0.098)	(0.122)	(0.098)	(0.115)	(0.276)	(0.093)
	[-0.447,0.031]	[-0.970,0.215]	[-0.139,0.307]	[-0.424,0.161]	[-0.288,0.193]	[-0.13,0.396]	[-1.123,0.11]	[-0.293,0.133]
Panel C: Excluding Union Decertification Elections								
Share newly unionized	-0.436	-0.444	-0.091	-0.417	-0.260	0.341	-0.626	-0.188
	(0.174)	(0.449)	(0.180)	(0.203)	(0.163)	(0.208)	(0.480)	(0.154)
	[-0.857,-0.046]	[-1.461,0.502]	[-0.523,0.212]	[-0.847,0.118]	[-0.627,0.169]	[-0.381,0.469]	[-2.470,-0.131]	[-0.581,0.139]
Panel D: State, Year, Industry Fixed Effects								
Share newly unionized	-0.223	-0.316	0.029	-0.189	0.006	0.277	-0.629	-0.091
	(0.205)	(0.335)	(0.209)	(0.223)	(0.207)	(0.218)	(0.376)	(0.197)
	[-0.76,0.232]	[-1.011,0.565]	[-0.504,0.497]	[-0.76,0.318]	[-0.526,0.475]	[-0.279,0.727]	[-1.492,0.307]	[-0.632,0.316]
Panel E: No Fixed Effects								
Share newly unionized	-0.323	-0.434	0.034	-0.242	-0.079	0.244	-0.543	-0.145
	(0.312)	(0.382)	(0.224)	(0.336)	(0.318)	(0.293)	(0.440)	(0.222)
	[-1.085,0.392]	[-1.266,0.561]	[-0.506,0.531]	[-1.058,0.538]	[-0.831,0.643]	[-0.449,0.865]	[-1.541,0.536]	[-0.723,0.331]
Panel F: Unweighted								
Share newly unionized	-0.384	-0.352	-0.193	-0.283	0.087	0.236	-0.848	-0.137
	(0.163)	(0.365)	(0.181)	(0.185)	(0.130)	(0.151)	(0.549)	(0.132)
	[-1.057,0.012]	[-2.102,0.583]	[-0.747,0.278]	[-0.817,0.407]	[-0.629,0.357]	[-0.412,0.691]	[-2.323,0.377]	[-0.828,0.086]

Notes: This table performs the robustness checks of Table A6 for the outcomes of Table 5. Please refer to those tables for the specifications and the data construction.
Sources: Decennial Census and NLRB unionization data; authors' calculations.

Table A9: Definitions of Decades

Decade	NLRB	Decennial Census and ACS	CPS MORG (for union density)	CPS ASEC (for fringe benefits)
	(1)	(2)	(3)	(4)
1960s	1961–69	1970 minus 1960	n/a	n/a
1970s	1970–79	1980 minus 1970	n/a	n/a
1980s	1980–89	1990 minus 1980	1990–1994 minus 1977–1980[b]	1990–1994 minus 1979–1981[b]
1990s	1990–99	2000 minus 1990	2000–2004 minus 1986–1990	2000–2004 minus 1986–1990
2000s	2000–09[a]	2009–2013 minus 2000	2010–2014 minus 1996–2000	2010–2014 minus 1996–2000

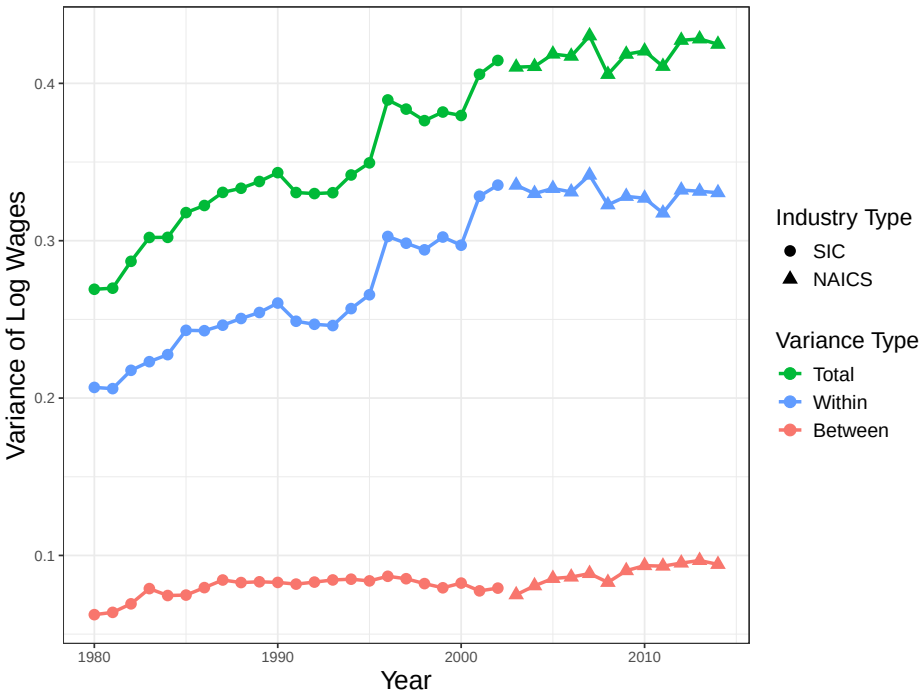
Notes: For each decade in the data, column 1 reports the years of NLRB data used to compute the new unionization rate. Columns 2–4 report the years in the decennial census (plus ACS) data, CPS MORG, and CPS ASEC data used to compute differenced outcomes. For example, the change over 1990s in the CPS MORG is computed as the average of 2000–2004 minus the average of 1986–1990.

[a] The NLRB data for the year of 2009 cover only the beginning of year.

[b] The May supplement to the CPS replaces CPS MORG for years 1977–1980 and replaces CPS ASEC for the year of 1979.

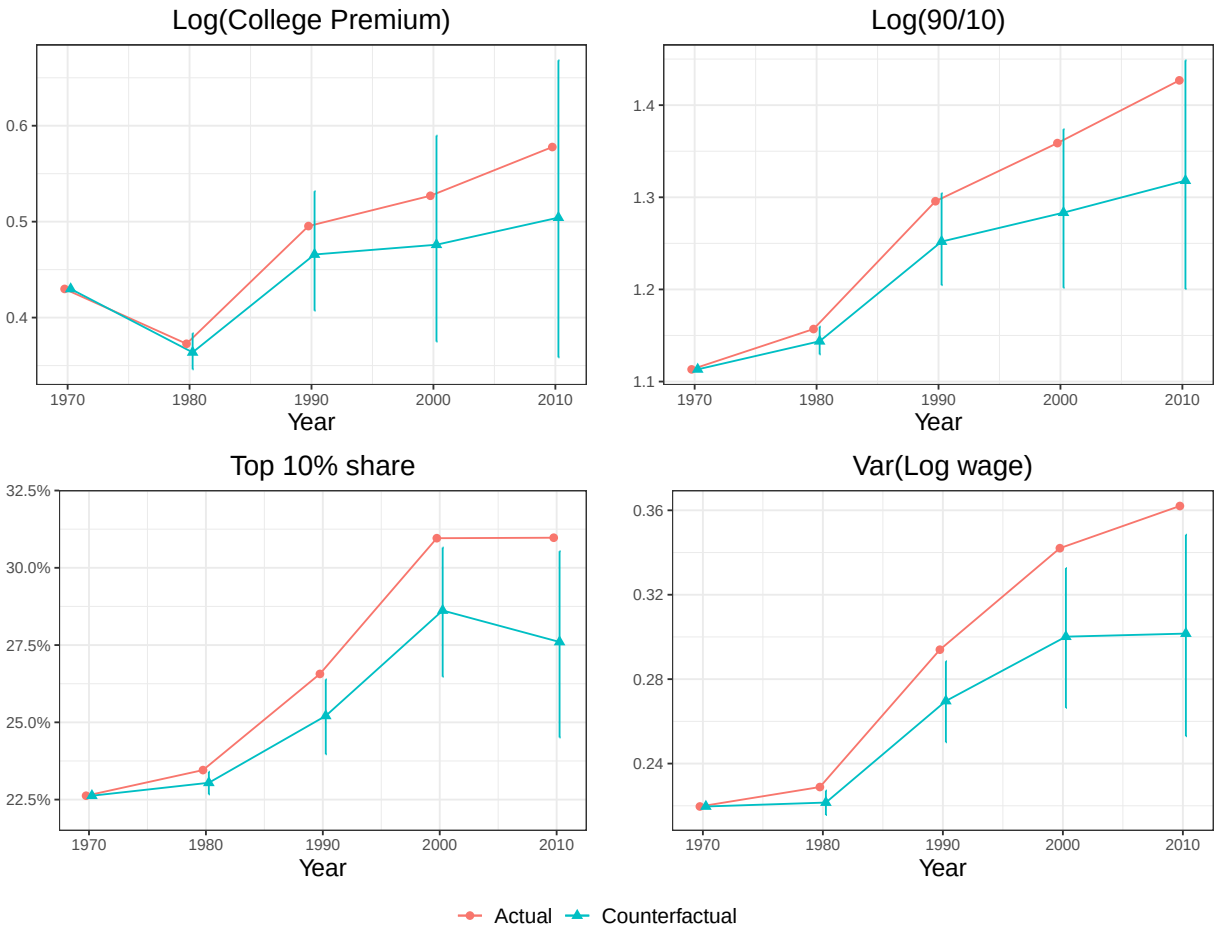
Appendix Figures

Figure A1: Within-Between Variance Decomposition of Log Wages, CPS Data



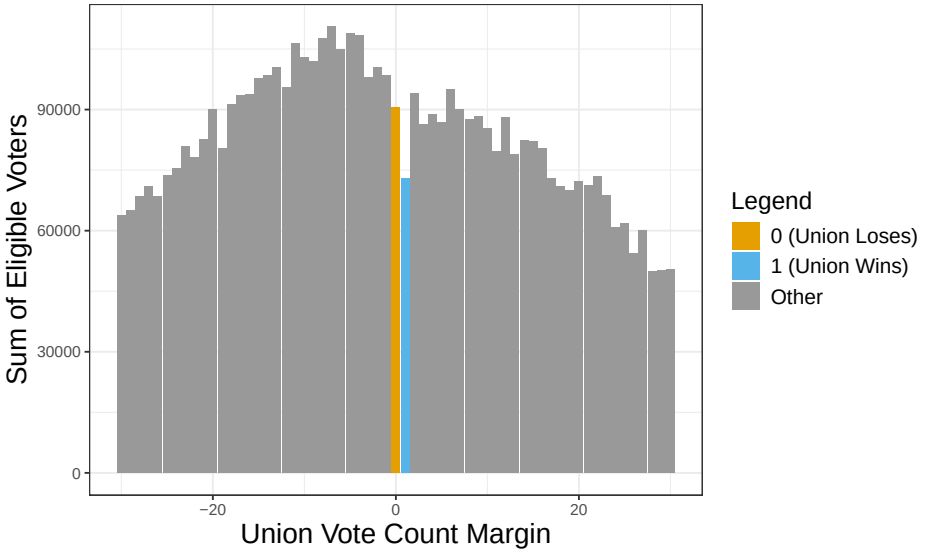
Notes: The plot presents a variance decomposition of log wages, broadly following Helpman et al. (2016), using CPS ASEC samples. In each year from 1980–2014, we decompose the overall variance in log wages among workers into within and between (state-industry) cell components, using projection weights. The industry coding framework shifts from SIC-based to NAICS-based in 2003.
Sources: CPS ASEC samples; authors' calculations.

Figure A2: The Counterfactual with Unionization Remaining at the 1960s Level: Other Inequality Outcomes



Notes: Each panel presents a measure of actual within (state-industry) cell inequality (averaged across cells with beginning-of-decade employment weights) by census year from 1970–2010, as well as its counterfactual value had the rate of new unionization stayed at the 1960s level. The counterfactual is calculated by adding to the actual measure the product of the treatment effect estimated in Table 3 and the cumulative shortfall in the rate of new unionization between the actual and counterfactual scenarios (presented in the first panel of Figure 3). Confidence intervals for the counterfactuals are based on the 95% robust bias-correct confidence intervals for the corresponding treatment effects.
Sources: Decennial Census, NLRB unionization data; authors' calculations.

Figure A3: Distribution of Total Eligible Voters by Union Vote Count Margin



Notes: For each margin between -30 and 30 votes, the histogram shows the total number of eligible workers in workplaces where unionization elections ended with this vote margin (i.e., the gap between the votes in favor and against the union). Tied elections and elections in which the union wins by one vote are highlighted. The time rage is 1961-2009.
Sources: NLRB unionization data; authors' calculations.