Regression Discontinuity Design

The Treatment Model

$W_i = 1$: Treatment (union, financial aid)
$W_i = 0$: Control (nonunion, non-financial aid)
$Y_i(0)$: Outcome if not treated
$Y_i(1)$: Outcome if treated

$$Y_i = (1 - W_i) \times Y_i(0) + W_i \times Y_i(1) = \begin{cases} Yi(0) & \text{if } W_i = 0 \\ Y_i(1) & \text{if } W_i = 1 \end{cases}$$

Sharp Regression Discontinuity Design

The treatment status is determined as follows

$$W_i = 1\{X_i \geq c\}$$

X_i: variable determining treatment (vote share, college financial aid score)
c : threshold. (50% vote share, aid score threshold)
Estimate the average treatment effect at the threshold point c

$$\tau_{SRD} = E[Y_i(1) - Y_i(0)|X_i = c]$$

Assumptions

Assumption 1

$$E\left[Y(0)|X=x\right],\ E\left[Y(1)|X=x\right]$$

are continuous in x.

Assumption 2

$$lim_{x\nearrow c}W_i(x)=0,\ lim_{x\searrow c}W_i(x)=1$$

Sharp RD Design

Then, we can use the following Sharp Regression Discontinuity Design Estimator to estimate the treatment effect.

$$\tau = lim_{x\searrow c}E\left[Y_i|X=x\right] - lim_{x\nearrow c}E\left[Y_i|X=x\right]$$

Economic Impacts of New Unionization on Private Sector Employers: 1984-2001

Lee and DiNardo

2004

Union Ballot

- Estimate the effect of unions on private sector plant survival, employment, wages, output, productivity.
- The problem is that unionization could be an endogenous variable determined by factors that also affect plant survival, employment, wages, etc.
- Exploit the secret ballot outcome that determine unionization. There is randomness of the final voting outcome. They look at plants that were not unionized due to votes being barely below 50% and those that were unionized with votes barely above 50%.
- The authors assume any variation of votes around the 50% unionization threshold is random, thus among the plants with around 50% votes, unionization outcome can be considered random.

Data

- National Labor Relations Borad (NLRB): 1984-1999. Data on elections.
- Federal Mediation and Consiliation Service (FMCS): 1984-2001. Data on Union Contract Expiration. Measure of collective bargaining activity after the union election.
- InfoUSA: 2001. Data on subsequent business survival, employment output.
- Longitudinal Research Database (LRD) of U.S. Census:1974-2001. Employment, wages, output, investment, and productivity in the manufacturing sector.

Process of Unionization

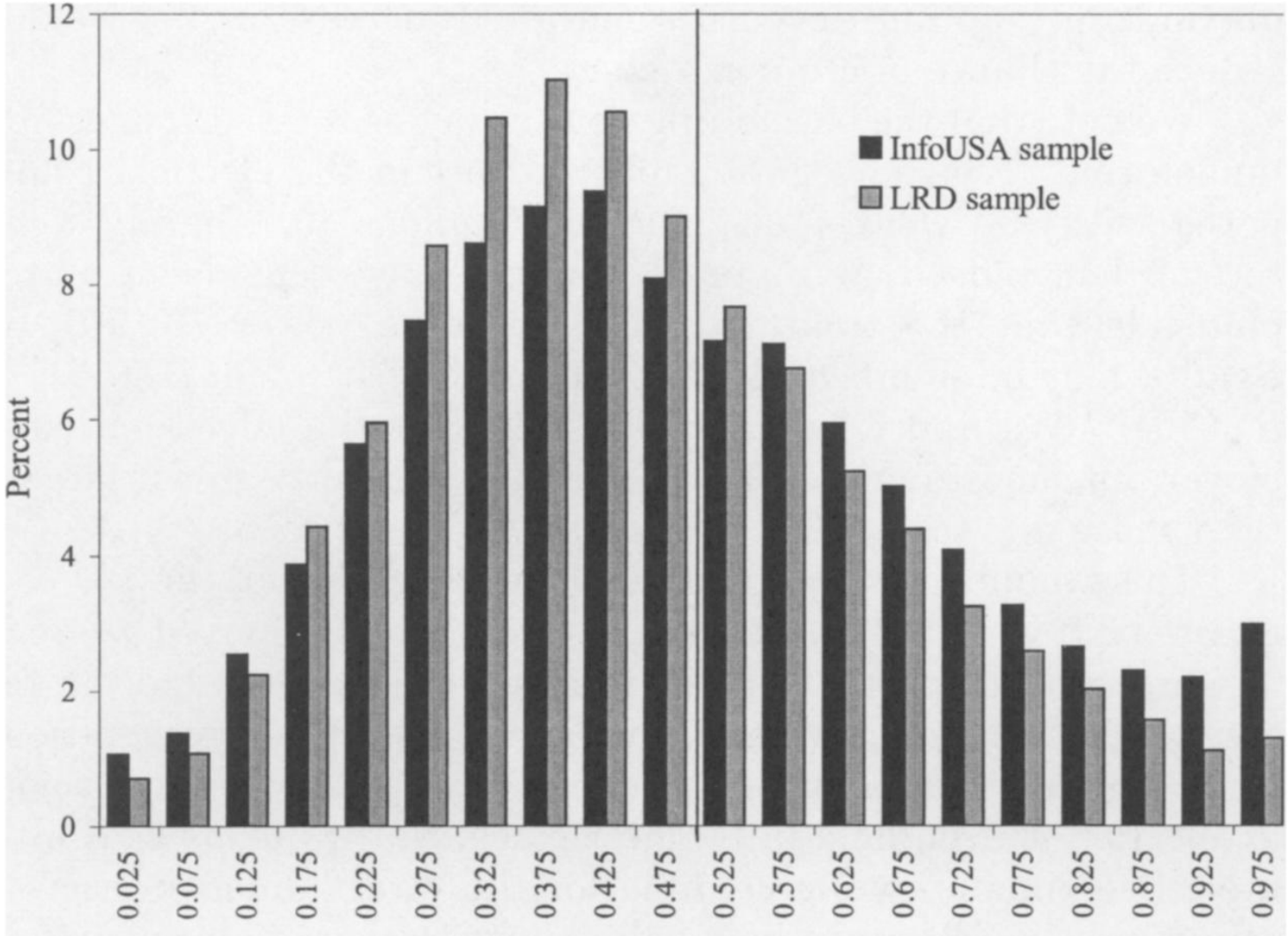
- A group or workers start organization drive with the assistance of a labor union.
- Card Drive: 50% approval cards are generally required (minimum 30%) in six months for a petition at the Natinal Labor Relations Board.
- The NLRB decides on the composition of workers in a bargaining unit.
- NLRB conducts the secret ballot election at the work site. 50% plus one or more votes are required for unionization. After hearing objections to elections, NLRB decides whether the results are valid or reelection needs to be done.
- If the union votes have majority, then the management has to negotiate in good faith with the union.

Firms, who often want to prevent unionization, can do the following.

- Hold meetings to discuss costs of unionization for the workers.
- Fire union activists without much punishments (only reinstatements).
- Hire anti-union consultants.
- Alleging unfaire labor practices, dispute the choice of bargaining unit.

Distribution of Union Vote Shares: All Certification Elections

50% threshold is not far away from the mode of the distribution.
Close elections are the norm of unioniation drives.


12
10
8
6
4
2
0
Percent
InfoUSA sample
LRD sample
0.025
0.075
0.125
0.175
0.225
0.275
0.325
0.375
0.425
0.475
0.525
0.575
0.625
0.675
0.725
0.775
0.825
0.875
0.925
0.975

Model

$$
\begin{aligned}
y &= X\gamma + D\beta + \epsilon \\
D &= 1(V > \frac{1}{2}) \\
V &= X\delta + u
\end{aligned}
$$

V: union vote: more than $\frac{1}{2}$ results in uninization.
β: measures the effect of unionization on the outcome y.
X in the equation determining V is the same as X in y. In this setup, there is no instrument to correct for endogeneity bias.

Sharp Regression Discontinuity Design

Assumptions

- There is some ex ante uncertainty in the vote share.
- The density of u conditional on X and ϵ is continuous.

Then, one can adopt the Sharp Regression Discontinuity Design and

$$\beta = lim_{\Delta \searrow 0^+} E\left[y|V = \frac{1}{2} + \Delta\right] - lim_{\Delta \nearrow 0^-} E\left[y|V = \frac{1}{2} - \Delta\right]$$

- Without uncertainty in vote share, $E[Y(0)|V = v]$ and $E[Y(1)|V = v]$ may not be continuous at the threshold point $\frac{1}{2}$. Only when there is a strong benefit to unionization, would union members push the vote over the 50%. Otherwise, the vote would not make it above 50%, implying discontinuous union benefits at 50% vote share.
- If the density of u conditional on X and/or ϵ is not continuous. Then, the distribution of X and/or ϵ conditional on $V = X\delta + u$ can be discontinuous at $\frac{1}{2}$, which makes $E[Y(0)|V = v]$ and $E[Y(1)|V = v]$ discontinuous at $v = \frac{1}{2}$.

- Continuity means that the density of X and ϵ will be approximately the same between treatment (unionized) and control (nonunionized) groups within the small neighborhood of $v = \frac{1}{2}$.
- That is, one can test the validity of RD design by looking at the $E[X|v]$ and see whether it is continuous around $V = \frac{1}{2}$.
- If not, then RD design is not valid because one does not know whether discontinuity in outcome comes from unionization or from the discontinuity in X.

Estimating Equation

$$Y_i = \alpha + \beta 1(V_i \geq \frac{1}{2}) + \sum_k \gamma_k V_i^k + \omega_i$$

Results

Union certification and bargaining

- Barely winning and barely losing an election have immediate and lasting impact of unionization. The fraction immediately recognized is essentially zero below the 50 percent threshold, and essentially 100 percent above the threshold.
- Close losers have much lower probability of eventually being unionized than close winners. Eventual desertification of close winners are also rare.
- Discontinuity of collective bargaining probability at 50% of vote count.

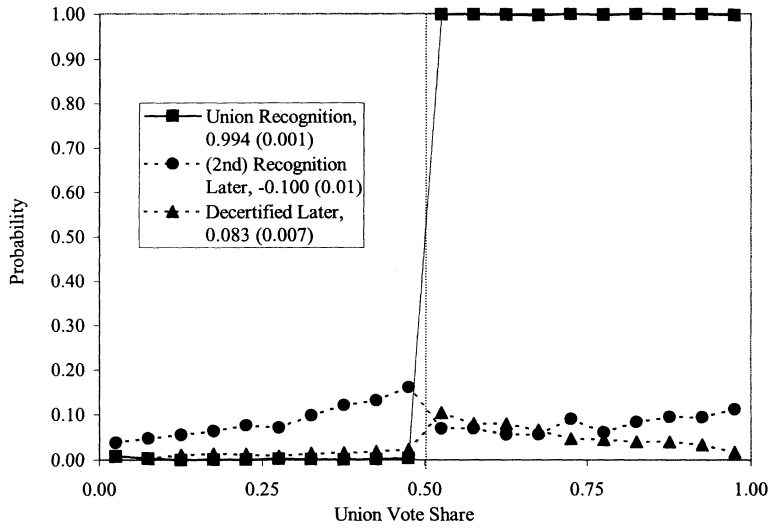


FIGURE IIIa
Recognition, Subsequent Certification or Decertification, by Union Vote Share.

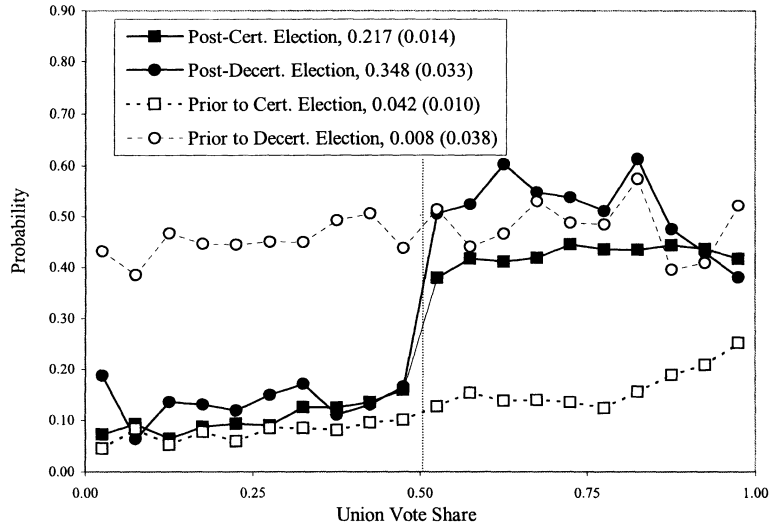


FIGURE IIIb
Contract Expiration Notice Filed, Prior to and Postcertification or Decertification Election, by Union Vote Share

Note: Figure IIIa: Initial Elections that take place between 1984–1995, 21405 observations. Point estimates and standard errors (in parentheses) are from a regression of the dependent variable on a fourth-order polynomial and a certification status dummy variable. Figure IIIb: Post-: Elections take place (1984–

Log Employment and Survival

- There is no discontinuity of log employment and firm survival in 2001 at 50% vote share. That is, no sample selection bias for the Regression Discontinuity Design.
- No discontinuity of log sales and log sales per worker.
- Pre-election bargaining probability, log hours, log output per hour, log wage: no discontinuity at 50%: supports randomness of votes around discontinuity.
- The regression estimates of the discontinuity could be sensitive to the functional form used, i.e. number of polynomials, etc.

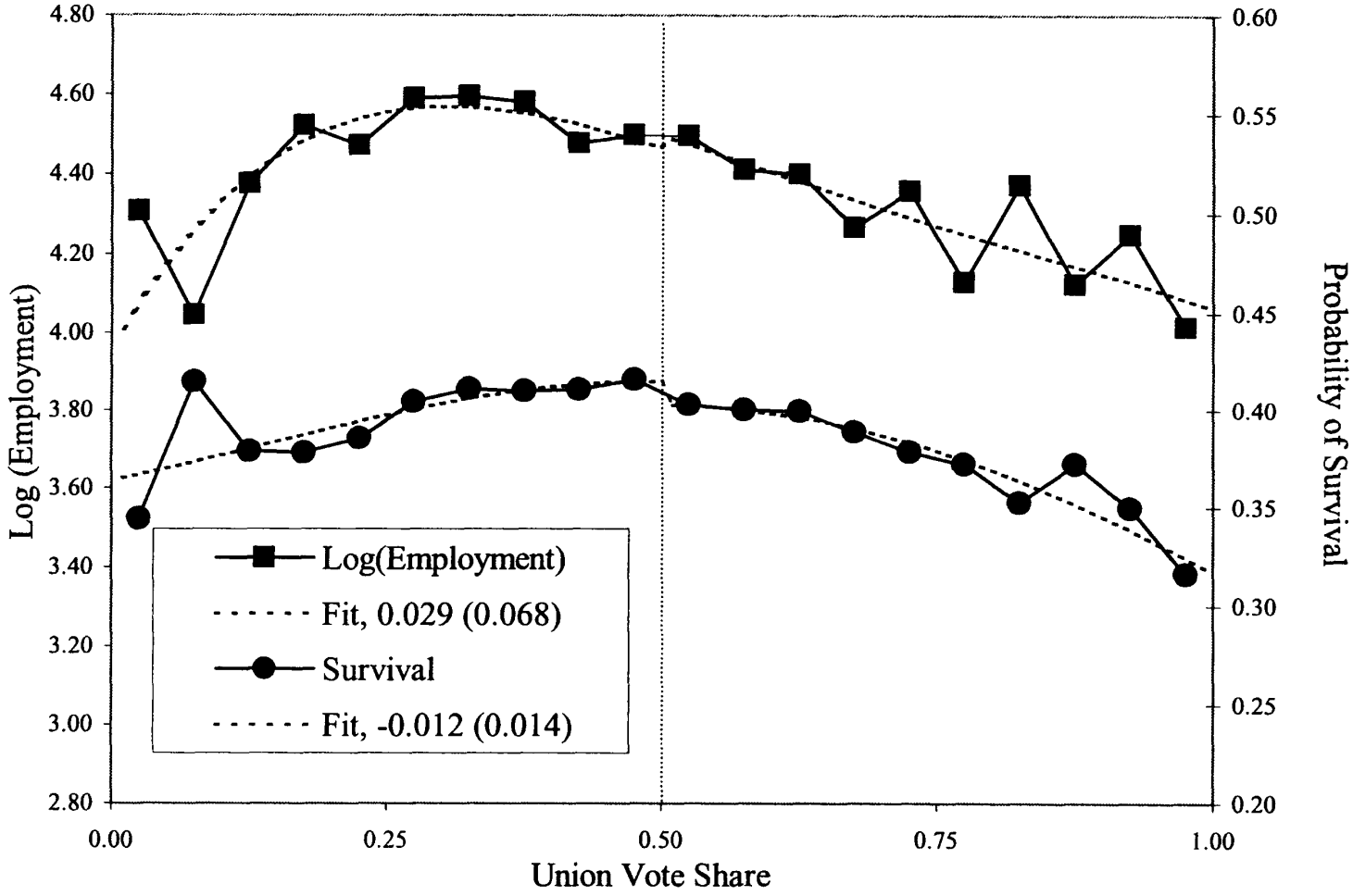


FIGURE IV

Log(Employment) and Survival at Year 2001, by Union Vote Share

TABLE I

OLS AND REGRESSION-DISCONTINUITY ESTIMATES, IMPACT OF UNION RECOGNITION ON BUSINESS SURVIVAL, LRD SAMPLE

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Union won	−0.089	−0.092	−0.078	−0.073	−0.046	−0.041	−0.027	−0.026	−0.021	−0.025	−0.021
	(0.013)	(0.013)	(0.013)	(0.013)	(0.022)	(0.023)	(0.029)	(0.029)	(0.028)	(0.028)	(0.028)
Vote share	—	—	—	—	−0.130	0.182	0.499	0.379	0.552	0.120	−0.062
					(0.055)	(0.134)	(0.398)	(0.732)	(0.720)	(0.721)	(0.730)
(Vote share)^2	—	—	—	—	—	−0.322	−1.110	−0.630	−1.405	−0.533	0.143
						(0.130)	(0.946)	(2.670)	(2.618)	(2.615)	(2.638)
(Vote share)^3	—	—	—	—	—	—	0.516	−0.210	0.862	0.415	−0.551
							(0.617)	(3.886)	(3.804)	(3.797)	(3.818)
(Vote share)^4	—	—	—	—	—	—	—	0.365	−0.106	−0.133	0.331
								(1.947)	(1.905)	(1.902)	(1.907)
Log(Votes cast)	—	—	0.061	0.065	—	—	—	—	—	0.059	0.064
			(0.008)	(0.008)						(0.008)	(0.008)
Year dummies?	No	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Yes
Industry dummies?	No	No	No	Yes	No	No	No	No	No	No	Yes
R^2	0.0084	0.0614	0.0729	0.0893	0.0094	0.0106	0.0107	0.0107	0.0637	0.0743	0.0907

Sample size is 5608. Some elections match to same establishment, so standard errors are clustered on the establishment level (4816 independent establishments). Includes all elections that occurred before 1997 that has the same name and address as a preelection-year entry in the LRD. Dependent variable is whether establishment appears in the 1997 Census of Manufactures. Least squares point estimates are reported.