

Online Appendices are the responsibility of the authors;
the material is not copyedited or formatted by the *ILR Review*.

Supplemental Online Appendix

Labor Unions and Workplace Safety

Ling Li, Shawn Rohlin, and Perry Singleton

Appendix

Data and Sample

The empirical analysis utilizes data from multiple sources. Data on union elections come from the National Labor Relations Board (NLRB) and the American Federation of Labor and the Congress of Industrial Organizations (AFL-CIO); data on union activity come from the Federal Mediation and Conciliation Service (FMCS); and data on occupational safety come from the Occupational Safety and Health Administration (OSHA), specifically the OSHA Data Initiative (ODI). These data are matched at the establishment level.

The NLRB reports the establishment name, address, and industry, as well as the number of eligible voters, valid votes cast, and votes for and against unionization. The NLRB data are first restricted to closed cases. To match the NLRB data to the other data, the establishment name and address were standardized. For the establishment name, all the special characters and common words, such as company, limited, and corporation, were deleted. If the listed formal name and the case name differed, or if the establishment is doing business as (DBA) under a different name, both names are retained and used for matching. For the street address, all special characters and numbers for floor, suite, and room were deleted. Common words, such as street, avenue, and road, were replaced with their respective abbreviations. To standardize and clean the city name, each name was best matched probabilistically to an exhaustive list of all city names in the US, compiled by the US Census Bureau. City names without a perfect match were checked manually for typos.

The NLRB were matched to the ODI and the FMCS. The ODI reports accident case rates, measured annually per 100 full-time equivalent workers. The FMCS data indicate whether an establishment filed a "notice of bargaining," an indicator of union activity. The establishment name and street address from the FMCS data and the ODI data were standardized using the same method of standardizing NLRB data.

The matching procedure utilized the establishment name, street address, city, state, and zip code. The NLRB was matched to the FMCS and ODI in several stages. In the first and most restrictive stage, the data were matched exactly on the establishment name, street, city, and state. If a record contained multiple establishment names, the matching procedure was repeated for each name until a successful match, if any. In the second stage, the data were matched exactly on the establishment name, zip code, city, and state. In the third stage, the data were matched exactly based on the first six letters of the establishment name and address within the same city and state. In the fourth and least restrictive stage, the data were matched exactly on the first six digits of the establishment name, the first six digits of the street, full city, and state. If an FMCS or ODI record successfully matched in one stage, the matched record was removed from matching in subsequent stages.

Support for Identification Strategy

Union Activity

The empirical strategy assumes that bargaining power of workers increases discontinuously at the 50-percent cutoff following the union election. To support this assumption, the election data from 1999 to 2010 are matched to “notices of bargaining” data in years 1997 to 2016 from the Federal Mediation and Conciliation Service (FMCS).¹ A notice is required to initiate, terminate, or modify a labor contract and is therefore an indicator of union activity. The NLRB elections are matched to FMCS records by establishment name and address.

Using the FMCS data on notices of bargaining, Appendix Figure 1 plots the FMCS match percent by calendar year before and after the union election, separately by the election outcome. Given the calendar years of the FMCS data, the sample is restricted to elections that occurred in 1999 and after, and the FMCS match rate only reflects periods for which data are available. Among establishments in which the unionization won, the match rate increases sharply in the calendar year of the election and the year after, then returns to its pre-existing trend. Among establishments in which unionization failed, the match rate remains relatively unchanged compared to the pre-existing trend. These results confirm that union activity increases following a successful election.

To examine union activity at the cutoff, Appendix Figure 2 plots the FMCS match rate in periods 0 or 1 across 20 non-overlapping bins of five percentage point each. Periods 0 and 1 had the greatest increase in union activity following a successful union election (Appendix Figure 1), and the number of bins equals the minimum votes in the sample so that establishments with the minimum votes can be represented in each bin (DiNardo and Lee, 2004). Given the calendar years of the FMCS data, the sample is restricted to elections that occurred from 1999 to 2009. The figure confirms that union activity increases

¹DiNardo and Lee (2004) similarly match union election data to FMCS data.

discontinuously at the cutoff following a union election.

The first row of Appendix Table 1 presents discontinuity estimates in the FMCS match rate in periods 0 or 1 using equation (1). Each figure corresponds to a discontinuity estimate from a single model. The rows correspond to outcome variables, and the columns correspond to different bandwidths and polynomial orders, which range from 0.15 to 0.25 and from 1 to 2, respectively. The first row presents discontinuity estimates for the FMCS match rate. As shown, the match rate increases discontinuously by more than 30 percentage points, which is statistically significant at the one percent level and robust to the bandwidth and polynomial order.

The FMCS match rates raise two issues. First, the FMCS match rate is not 100 percent among establishments with a successful union election. One possible reason is that the match of the NLRB data to the FMCS data is imperfect, leading to false-negatives. Another possible reason is that a successful union election may not necessarily lead to a labor contract. Second, the FMCS match rate is not zero among establishments with a failed union election. Again, one possible reason is that the match of the NLRB data to the FMCS data is imperfect, although false-positives seem less likely than false-negatives. Another reason is that a single establishment may have multiple bargaining units, and the bargaining unit associated with the union election may be different from the bargaining unit associated with a FMCS match.

Density of the Vote Share

A threat to identification is non-random sorting at the cutoff, which occurs when the vote share is manipulated at the margin of victory to alter the election outcome. The problem for identification is that manipulation for or against a successful election may correlated with workplace safety so that the discontinuity in workplace safety at the cutoff may be biased relative to the true causal effect of unionization.

Non-random sorting due to vote share manipulation may lead to excess or missing

density at the cutoff, so a discontinuity in the density of the vote share may be evidence of non-random sorting (McCrary, 2008). To illustrate the vote share distribution, the top panel of Appendix Figure 3 plots the vote share density distribution across 20 non-overlapping bins of five percentage point each. The figure indeed suggests manipulation: the density generally increases from the right towards the cutoff, but decreases just above the cutoff, suggesting too few narrow election successes. This finding is consistent with an earlier study by Frandsen (2014), who also finds too few narrow election successes.²

A test for continuity at the cutoff is developed by McCrary (2008), which is based on the local linear density estimator developed by Cheng (1994) and Cheng et al. (1993). Using a uniform kernel, the test fails to reject continuity at the 5 percent level (p-value: 0.0696), though the test statistic is only marginally insignificant.

Manipulating the vote share may become increasingly more difficult as the number of votes increases. The second and third panels of Appendix Figure 3 illustrate vote share densities among elections with at least 50 and 100 votes, respectively. Although both figures indicate missing density just above the cutoff, the McCrary (2008) test fails to reject continuity. The failure to reject may be partially due to a smaller sample, which is reduced by approximately 50 percent and 75 percent when examining elections with at least 50 votes and 100 votes, respectively.

There are two potential strategies to address excess or missing density at the cutoff. The first is to focus on union elections with many votes cast, where vote share manipulation to affect the election outcome is not only more difficult, but more likely to be unexpected locally at the cutoff (Lee, 2008). As McCrary (2008) notes, excess or missing density does not necessarily invalidate identification if manipulation occurs at random. The second strategy is to evaluate non-random sorting directly with respect to observable characteristics and workplace safety pre-election.

²A key difference between between the studies is that Frandsen (2014) uses elections from 1980 to 2009, whereas this study uses elections from 1991 to 2010. This study also uses only establishments with union elections that match to the ODI, which are not representative of all establishments with union elections.

Establishment and Election Characteristics

Non-random sorting may also lead to discontinuities in the conditional mean functions of observable characteristics (Lee, 2008). To evaluate discontinuities in election characteristics graphically, the first row of Appendix Figure 4 plots the conditional mean of eligible employees and valid votes cast across 20 mutually exclusive bins.³ As shown, the number of eligible employees and valid votes cast increases and then decreases with the vote share, but there is no apparent discontinuity in either measure at the cutoff.

The second and third rows of Appendix Table 1 present discontinuity estimates in eligible employees and valid votes using equation (1). As shown, the discontinuity estimates are generally small in magnitude relative to the mean near the cutoff, and all of the estimates are statistically insignificant. The largest estimates appear in the second column, with a bandwidth of 0.20 and a second-order polynomial. In that column, the discontinuity estimates are -5.83 and -3.07 for eligible employees and valid votes, respectively, relative to the mean near the cutoff of 116.98 and 102.90.⁴

To evaluate discontinuities in industry, the second row of Appendix Figure 4 plots the share of establishments in manufacturing and health services across 20 mutually exclusive bins. As shown, the share in manufacturing increases and then decreases with the vote share, whereas the share in health services generally increases, except for a sharp decrease at the highest bins. Importantly, neither industry share exhibits a discontinuity at the cutoff.

The fourth and fifth rows of Appendix Table 1 presents discontinuity estimates in the industry share using equation (1). In this case, all the estimates are small and statistically insignificant regardless of the bandwidth or polynomial order. These results suggest that there is no non-random sorting around the cutoff with respect to industry.

Finally, the last panel of Appendix Figure 4 plots the share of establishments with

³When examining discontinuities in the number of eligible employees and votes cast, one extreme outlier is omitted from the analysis. For this outlier, the number of eligible employees and votes case is 17,195 and 15,471, respectively. The next highest values are 7,000 and 4,589, respectively.

⁴The means near the cutoff are calculated among elections within a bandwidth of 10 percentage points from the cutoff.

any match to the ODI.

APPENDIX TABLE 1: DISCONTINUITY ESTIMATES IN UNION ACTIVITY AND ESTABLISHMENT AND ELECTION CHARACTERISTICS

Outcome Variable	Mean	(1)	(2)	(3)
FMCS	21.63 (0.70)	35.00*** (2.54)	32.47*** (3.40)	32.97*** (3.01)
Employees eligible (number)	116.98 (1.92)	2.412 (6.733)	-5.834 (9.099)	-1.360 (7.883)
Valid votes (number)	102.90 (1.64)	3.059 (5.626)	-3.070 (7.633)	1.535 (6.600)
Manufacturing	32.39 (0.54)	-0.802 (1.779)	0.807 (2.375)	-0.682 (2.100)
Health services	20.30 (0.46)	0.884 (1.582)	0.261 (2.117)	1.127 (1.865)
ODI Match, Periods -3 to -1	18.07 (0.56)	-0.423 (1.888)	-0.870 (2.523)	0.947 (2.226)
ODI Match, Periods 1 to 3	19.22 (0.48)	1.25 (1.621)	0.640 (2.174)	1.89 (1.913)
Polynomial		1	2	2
Bandwidth		0.15	0.20	0.25
Observations		11,150	14,374	17,187

The sample is derived from union elections contained in the National Labor Relations Board data, file years 1991 to 2010, and is restricted to elections with at least 20 valid votes and a valid vote share in favor of unionization. Each column corresponds to discontinuity estimates for one dependent variable. The FMCS is an indicator for union activity. Due to the dates of the FMCS data, the sample for these regressions is restricted to elections in 1999 to 2009. For employees eligible and valid votes, one outlier is dropped from the sample. The mean is calculated with a bandwidth of 10 percentage points above and below the cutoff. Standard errors are in parentheses and clustered by election. Estimates are in percentage points unless otherwise noted. FMCS, Federal Mediation and Conciliation Service; ODI, OSHA Data Initiative.

***, **, and * indicate significance at the one, five, and ten percent levels, respectively.

APPENDIX TABLE 2: DISCONTINUITIES IN DART POST-ELECTION (PERIODS 1 TO 3),
 $20 \leq \text{VOTES} < 70$

	Mean	(1)	(2)	(3)	(4)	(5)	(6)
DART (rate)	7.93 (0.22)	0.0247 (0.876)	0.0256 (1.031)	-0.563 (1.124)	-0.637 (1.291)	-0.106 (1.020)	0.102 (1.182)
DART=0	8.35 (0.83)	2.971 (3.344)	2.335 (3.582)	3.449 (4.819)	2.423 (5.088)	3.271 (4.159)	3.761 (4.320)
DART<1	10.32 (0.91)	1.024 (3.678)	0.569 (4.135)	1.038 (5.276)	-1.828 (5.821)	1.702 (4.535)	1.581 (4.937)
DART<2	16.70 (1.12)	1.404 (4.423)	0.547 (4.812)	0.768 (6.263)	-1.929 (6.605)	2.228 (5.396)	0.926 (5.734)
DART<3	25.31 (1.30)	0.905 (5.331)	0.438 (5.710)	3.210 (7.450)	1.943 (7.906)	4.072 (6.428)	1.691 (6.728)
DART<4	32.32 (1.40)	-0.750 (5.816)	-1.555 (6.265)	1.066 (8.124)	-1.195 (8.500)	0.879 (7.033)	-0.970 (7.413)
DART<5	38.96 (1.46)	-0.499 (5.988)	0.482 (6.643)	-0.271 (8.359)	-0.342 (8.936)	0.480 (7.249)	-0.195 (7.790)
DART<6	47.04 (1.50)	-8.830 (5.979)	-8.783 (6.624)	-7.900 (8.299)	-10.60 (8.746)	-7.620 (7.213)	-8.524 (7.645)
DART<7	55.57 (1.49)	-1.793 (5.951)	-1.432 (6.671)	1.243 (8.135)	1.736 (8.652)	0.405 (7.129)	1.166 (7.658)
DART<8	61.49 (1.46)	-1.085 (5.766)	-0.399 (6.604)	-0.898 (7.740)	-0.116 (8.436)	1.913 (6.855)	2.459 (7.550)
DART<9	66.34 (1.42)	-0.127 (5.721)	-0.952 (6.719)	2.318 (7.608)	1.633 (8.417)	2.698 (6.766)	2.339 (7.568)
DART<10	70.47 (1.37)	-1.627 (5.563)	-1.836 (6.801)	2.160 (7.420)	2.484 (8.431)	0.308 (6.576)	0.335 (7.591)
Covariates		No	Yes	No	Yes	No	Yes
Polynomial		1	1	2	2	2	2
Bandwidth		0.15	0.15	0.20	0.20	0.25	0.25
Observations		1,586	1,586	1,990	1,990	2,385	2,385

The sample is derived from union elections contained in the National Labor Relations Board data, file years 1991 to 2010, and is restricted to elections with at least 20 valid votes and a valid vote share in favor of unionization. To examine effects among smaller bargaining units, the sample is further restricted to elections with valid votes greater than or equal to 20 and less than 70. Observations are establishment-by-ODI match. The mean is calculated with a bandwidth of 10 percentage points above and below the cutoff. Standard errors are in parentheses and clustered by election. Estimates are in percentage points unless otherwise noted. DART, days away from work, job restrictions, and job transfers; ODI, OSHA Data Initiative.

***, **, and * indicate significance at the one, five, and ten percent levels, respectively.

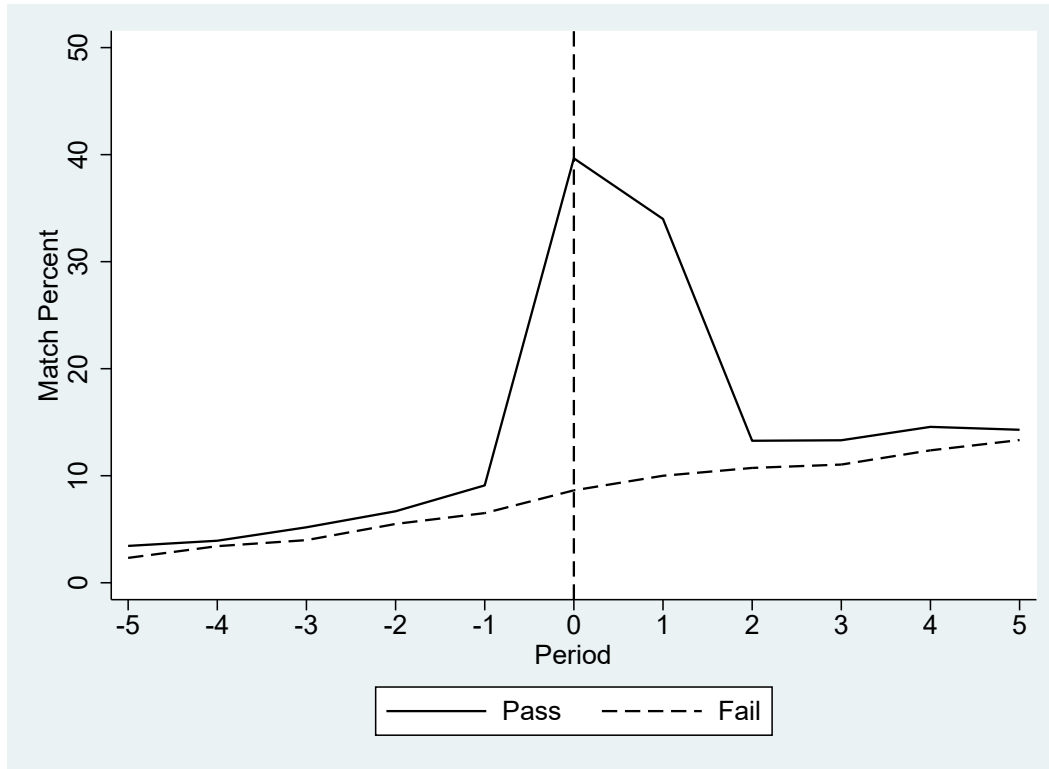
APPENDIX TABLE 3: DISCONTINUITIES IN DART POST-ELECTION (PERIODS 1 TO 3),
NOT MANUFACTURING

	Mean	(1)	(2)	(3)	(4)	(5)	(6)
DART (rate)	8.18 (0.22)	0.735 (0.935)	0.858 (1.037)	1.552 (1.224)	1.438 (1.299)	0.694 (1.126)	0.642 (1.188)
DART=0	5.35 (0.69)	2.084 (3.156)	3.346 (3.341)	3.242 (4.359)	3.446 (4.541)	2.569 (3.901)	2.886 (3.972)
DART<1	7.41 (0.80)	0.640 (3.513)	3.370 (3.881)	1.187 (4.868)	3.106 (5.337)	1.649 (4.332)	2.592 (4.568)
DART<2	12.85 (1.03)	1.406 (4.424)	6.241 (4.948)	0.611 (6.115)	5.863 (6.751)	1.201 (5.410)	4.338 (5.718)
DART<3	20.73 (1.24)	1.247 (5.101)	6.583 (5.918)	2.143 (6.969)	7.581 (7.914)	1.231 (6.156)	3.808 (6.707)
DART<4	27.67 (1.37)	-2.011 (5.630)	2.630 (6.740)	-2.447 (7.602)	2.173 (9.016)	-3.854 (6.753)	-1.573 (7.604)
DART<5	35.27 (1.46)	-4.254 (6.037)	-1.097 (7.065)	-8.622 (8.133)	-4.412 (9.426)	-6.237 (7.232)	-3.860 (7.897)
DART<6	43.71 (1.52)	-9.702 (6.259)	-9.307 (7.041)	-13.09 (8.390)	-13.01 (9.158)	-10.84 (7.479)	-12.11 (7.791)
DART<7	51.59 (1.53)	-2.203 (6.278)	-4.102 (6.902)	-5.348 (8.339)	-5.954 (8.887)	-2.983 (7.461)	-5.443 (7.639)
DART<8	57.04 (1.52)	-4.588 (6.188)	-6.342 (6.986)	-8.531 (8.082)	-8.910 (8.725)	-4.659 (7.304)	-6.225 (7.666)
DART<9	62.01 (1.49)	-3.937 (6.114)	-6.989 (7.018)	-6.783 (7.928)	-8.301 (8.698)	-4.105 (7.197)	-6.223 (7.615)
DART<10	67.92 (1.43)	-5.175 (5.906)	-7.240 (7.371)	-8.620 (7.656)	-10.12 (9.091)	-6.449 (6.942)	-7.912 (7.808)
Covariates		No	Yes	No	Yes	No	Yes
Polynomial		1	1	2	2	2	2
Bandwidth		0.15	0.15	0.20	0.20	0.25	0.25
Observations		1,539	1,539	1,923	1,923	2,274	2,274

The sample is derived from union elections contained in the National Labor Relations Board data, file years 1991 to 2010, and is restricted to elections with at least 20 valid votes and a valid vote share in favor of unionization. To examine effects by industry, the sample is further restricted to establishments not in manufacturing. Observations are establishment-by-ODI match. The mean is calculated with a bandwidth of 10 percentage points above and below the cutoff. Standard errors are in parentheses and clustered by election. Estimates are in percentage points unless otherwise noted. DART, days away from work, job restrictions, and job transfers; ODI, OSHA Data Initiative.

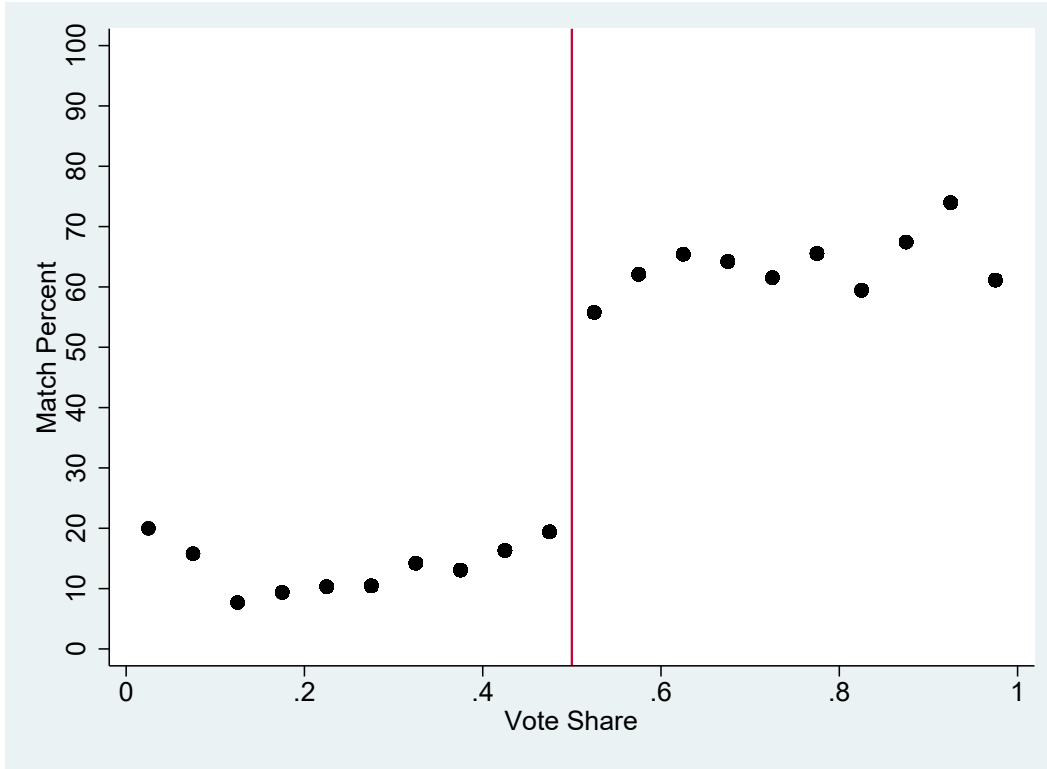
***, **, and * indicate significance at the one, five, and ten percent levels, respectively.

APPENDIX FIGURE 1: FMCS MATCH BY PERIOD



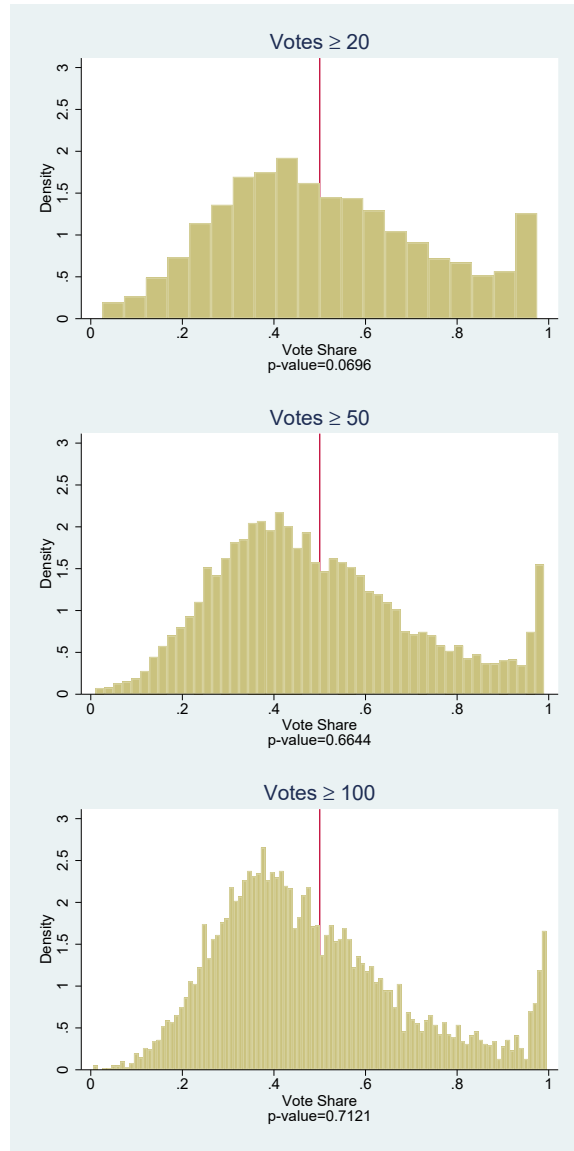
The sample is derived from union elections contained in the National Labor Relations Board data, file years 1999 to 2009, and is restricted to elections with at least 20 valid votes and a valid vote share in favor of unionization. The periods correspond to calendar years relative to the calendar year of the election. The match rates are calculated using only calendar years for which ODI data are available. ODI, OSHA Data Initiative.

APPENDIX FIGURE 2: FMCS MATCH BY VOTE SHARE



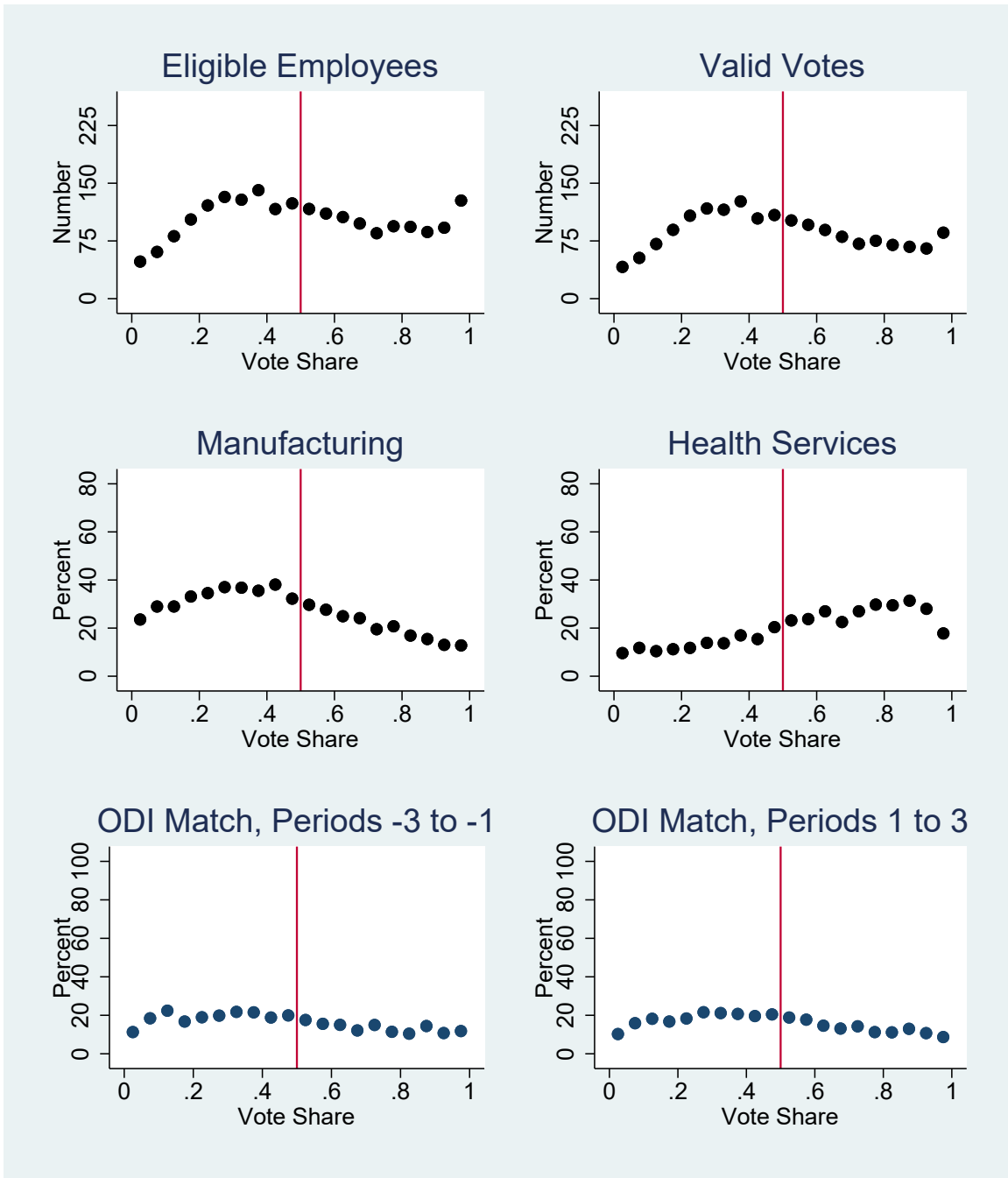
The figure illustrates the match rate of union elections contained in the National Labor Relations Board data to notices of bargaining filed with the Federal Mediation and Conciliation Service in periods zero or one across bins of five percentage points. The sample is derived from union elections contained in the NLRB, file years 1999 to 2009. The sample is restricted to elections with at least 20 valid votes and a valid vote share in favor of unionization.

APPENDIX FIGURE 3: DISTRIBUTION OF VOTE SHARE



The sample is derived from union elections contained in the National Labor Relations Board data, file years 1991 to 2010, and is restricted to elections with at least 20 valid votes and a valid vote share in favor of unionization. In each panel, the number of bins is equal to the minimum number of votes cast. To impose symmetry in the vote share distribution regardless of the number of vote cast, an amount equal to 0.5 divided by the number of votes cast is subtracted from the vote share if the number of votes cast is even (DiNardo and Lee, 2004). The p-value comes from the McCrary (2008) test for continuity at the cutoff.

APPENDIX FIGURE 4: DISCONTINUITIES IN ELIGIBLE EMPLOYEES, VALID VOTES CAST, INDUSTRY, AND ODI MATCH



The sample is derived from union elections contained in the National Labor Relations Board data, file years 1991 to 2010, and is restricted to elections with at least 20 valid votes and a valid vote share in favor of unionization.

References

- Cheng, Ming-Yen. 1994. On boundary effects of smooth curve estimators. Unpublished doctoral dissertation. University of North Carolina at Chapel Hill, Department of Statistics.
- Cheng, Ming-Yen, Jianqing Fan, and J.S. Marron. 1993. Minimax efficiency of local polynomial fit estimators at boundaries. Unpublished paper. University of North Carolina at Chapel Hill, Department of Statistics.
- DiNardo, John, and David S. Lee. 2004. Economic impacts of new unionization on private sector employers: 1984–2001. *Quarterly Journal of Economics* 119(4): 1383–442.
- Frandsen, Brigham R. 2014. The surprising impacts of unionization: Evidence from matched employer-employee data. Unpublished paper. Brigham Young University, Economics Department.
- Lee, David S. 2008. Randomized experiments from non-random selection in U.S. house races. *Journal of Econometrics* 142(2): 675–97.
- McCrary, Justin. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142(2): 698–714.