Labor Unions and Occupational Safety*

Ling Li† Shawn Rohlin ‡ Perry Singleton §

January 15, 2018

Abstract
This study attempts to identify the effect of unionization on occupational safety. Using
the regression discontinuity model, the empirical strategy uses establishments with elec-
tions that narrowly failed as a comparison group for establishments with elections that
narrowly passed. Data on elections come from the National Labor Relations Board,
and data on occupational safety come from the Occupational Safety and Health Ad-
ministration. According to the results, union activity increased sharply at the cutoff
for a successful union election - 50 percent plus one in favor of unionization - but ac-
cident case rates remained relatively constant.

Keywords: unions, occupational safety, OSHA
JEL Codes: J28, J51, J81

Corresponding Author: Perry Singleton; Syracuse University; 426 Eggers Hall;
Syracuse, NY 13244; psinglet@syr.edu

*For helpful comments and suggestions, the authors would like to thank Gary Engelhardt, Brigham
Frandsen, Hugo Jales, and Jeffrey Kubik. The authors would also like to thank Jeanette Walters-Marquez
for providing data from the Federal Mediation and Conciliation Service.
†Syracuse University, Department of Economics, and the Center for Policy Research
‡Kent State University, Department of Economics, and the Center for Entrepreneurship and Business
Innovation
§Syracuse University, Department of Economics, and the Center for Policy Research
1 Introduction

Workers form labor unions to bargain over wages, employment, and working conditions. While most research focuses on the determination of wages and employment (Farber, 1986), relatively little research focuses on working conditions. To address this limitation, this study examines the effect of unions on occupational safety. As Morantz (2013) notes, unions engage in numerous safety-enhancing activities, including pressuring employers to maintain safe workplaces, educating workers about workplace hazards, and developing safety-related innovations through economies of scale. The effect on occupational safety has direct implications for worker welfare and the efficiency of labor unions. The effect is also relevant to research on unions and wages (Branchflower and Bryson, 2004; DiNardo and Lee, 2004; Frandsen, 2014; Freeman and Medoff, 1984), since occupational safety may affect wages through compensating differentials (Kniesner and Leeth, 2014).

This study attempts to identify the effect of unionization on occupational safety. Following DiNardo and Lee (2004), Frandsen (2014) and others, the empirical strategy exploits the timing and outcome of union elections. Specifically, establishments in which elections narrowly failed are used as a comparison group for establishments in which elections narrowly passed. The difference in outcomes between the groups is estimated using a regression discontinuity model. Data on union elections come from the U.S. National Labor Relations Board (NLRB), and data on occupational safety come from the Occupational Safety and Health Administration (OSHA). These data contain accident case reported by employers at the establishment level. To examine union activity following an election, the data are matched to “notices of bargaining” filed with the Federal Mediation and Conciliation Service (FMCS). To examine firm survival after an election, the election data are matched to a national database of establishments compiled by InfoUSA. The analysis is limited to elections held in 1999 to 2010.

According to the results, union activity increased sharply at the cutoff for a successful union election - 50 percent plus one in favor of unionization - but there was relatively
no change in accident case rates. The discontinuity estimates are more precise for case rates involving days away from work, job restrictions, and job transfers. Using several different empirical specifications, the estimated effects of unionization on occupational safety are near zero, bounded below by the 95 percent confidence interval at approximately -1.5 cases annually per 100 full-time equivalent workers. The estimates are more precise when comparing all establishment just above and below the cutoff, not just at the discontinuity. For example, by comparing establishments within ten percent above the cutoff to establishments within ten percent below, the estimated difference in case rates is 0.177 with a confidence interval of -0.504 to 0.857. In the year of the election, the mean case rate was 7.99.

The study contributes to an existing literature on unionization and occupational safety. Despite the numerous safety-enhancing activities of unions, most empirical studies find that unionization is associated with greater accidents and injuries (Donado, 2015). One possible explanation is selection, whereby more dangerous establishments are more likely to unionize. A few studies find that unions improve occupational safety, but these findings pertain to specific eras and industries (Boal, 2009; Fairris, 1995; Morantz, 2013).

This study differs from related studies on several dimensions. First, the identification strategy exploits narrow union elections. While this strategy potentially addresses selection into unionization, the results pertain only to newly unionized establishments with narrow election victories. Second, the data are more recent and cover different industries, incorporating a large share of establishments in manufacturing and health services. Thus, while the identification strategy likely improves the internal validity of the results, the findings are less generalizable to other contexts or directly comparable to related studies.

A limitation of this study is that accident case rates are self-reported by the employer. The concern is that, while unionization may make workplaces safer, thereby decreasing case rates, unionization may also pressure employers to accurately record and report accident case rates, thereby increasing case rates. Thus, the negligible effect of unionization on case rates has two possible mechanisms: both occupational safety and reporting remain
unchanged, or occupational safety improves while reporting increases, with no change on net. To distinguish between the two mechanisms in future research, it is important to consider more objective measures of occupational safety.

2 Background

Workers form labor unions to create or capture monopoly rents resulting from imperfect competition in labor and product markets (Farber, 1986). A single union represents workers across multiple firms and establishments, forming national and increasingly international coalitions. At the establishment level, union officials represent workers during negotiations and help to enforce labor contracts. Labor contracts specify the terms and conditions of employment, such as wages, employment, and working conditions.

In the US, workers typically join unions through union elections.\(^1\) Elections are facilitated by the National Labor Relations Board (NLRB), established in 1935 to enforce collective bargaining laws. To hold an election, organizers must first demonstrate at least 30 percent support for a union election among eligible workers. This is achieved by petitions or authorization cards. If successful, the NLRB determines the size and scope of the bargaining unit and sets the time and location of the election. The election is conducted by secret ballot, and a successful election requires a simple majority. If an election is successful, employers must bargain “in good faith” with the union during contract negotiations.

A framework of union bargaining power is developed by DiNardo and Lee (2004). In their framework, bargaining power is a function of the share of workers who favor unionization. In the baseline case, where union elections are permitted, but none occur, bargaining power increases monotonically with the vote share. If workers successfully petition for an election, bargaining power increases further, independent of the election outcome. This is referred to as the indirect “threat” effect of an election. If the union election is successful,

\(^1\)In some cases, an employer will independently recognize a labor union, forgoing an election.
bargaining power increases even further. This is referred to as the direct effect of unionization. As DiNardo and Lee (2004) note, because a successful election requires a simple majority, bargaining power increases discontinuously at the 50-percent vote share, assuming the direct effect of unionization on bargaining power is non-zero.

In an analysis of union behavior, Farber (1986) considers two types of bargaining structures. In the first structure, unions bargain over wages only, leaving employers to determine employment levels and other conditions of employment. In this case, unions negotiate along the employer’s demand curve, increasing wages and decreasing employment, constrained by a non-negative-profit condition. However, the resulting labor contract is generally inefficient, as one party can be made better off without negatively impacting the other. In the second structure, unions bargain over both wages and employment. In this case, unions negotiate along the employer’s isoprofit curve, decreasing wages and increasing employment relative to the demand curve. In this case, the resulting labor contract may be efficient, but the net effect on wages and employment is ambiguous.

While most economic research on unionization has focused on wages and employment, unionization could impact other aspects of employment such as occupational safety. Indeed, unions engage in numerous safety-enhancing activities, including pressuring employers to maintain safe workplaces, educating workers about workplace hazards, developing safety-related innovations through economies of scale, and influencing the stringency of regulatory oversight (Morantz, 2009). In regards to regulatory oversight, unions lobbied for the establishment of the Occupational Safety and Health Administration (OSHA), which codifies and enforces occupational safety and health regulations (Mendeloff, 1980; Schurman et al., 1998). Unionized establishments also face more frequent and rigorous inspections under OSHA (Weil, 1991, 1992). In regards to education, unions increased awareness of the hazards from exposure to coal dust, cotton dust, asbestos, radium, and dibromochloropropane (Donado, 2015). Lastly, at the establishment level, unions work directly with management through safety and health committees, composed of workers, managers, and union officials
Despite the safety-enhancing activities of unions, the causal effect of unions on occupational safety is ambiguous. The conceptual framework is analogous to the discussion of wages and employment. On one hand, unions may bargain over wages only. In this case, employers may reduce investments in occupational safety as wage costs increase. Unions may also explicitly bargain for higher wages in exchange for decreased occupational safety. On the other hand, unions may bargain over both wages and occupational safety, potentially increasing both. The ability to contract on occupational safety depends, in part, on whether occupational safety can be monitored.

In the empirical literature, most studies find that unionization is associated with greater accidents and injuries (Donado, 2015). The literature considers three primary explanations. The first is selection, whereby establishments with greater accidents and injuries are more likely to unionize (Hills, 1985). The second is reporting, whereby employers are more likely to report accidents and injuries to federal agencies in the presence of union representation. This may be due to the employer’s tendency to underreport, the union’s tendency to overreport, or both. The third explanation, stated above, is that unions bid up wages in exchange for safety.

A few studies find positive effects of unions on occupational safety, but these findings pertain to specific eras and industries. For example, Boal (2009) examines turn-of-the-century coal mining, and Fairris (1995) examines company unions in the 1920s. One study by Morantz (2013) focuses on mining-related injuries and fatalities in the 1970s and 1980s. However, the variation in union status is mostly between mines, rather than within, as few coal mines transitioned between union and nonunion states. A particular concern is that dangerous mines are more likely to unionize, which would bias downward the estimated effect of unionization on occupational safety.
3 Empirical Strategy

The empirical objective is to identify the effect of unionization on occupational safety. A key issue for identification, as stated above, is that more dangerous firms may be more likely to unionize. To attempt to address this issue, this study pursues an empirical strategy by DiNardo and Lee (2004) and Frandsen (2014) and others that focuses on close union elections. If establishments with elections that narrowly fail are comparable to those with elections that narrowly pass, then the former may serve as a counterfactual for the latter.

The empirical strategy follows the regression discontinuity design (Hahn et al., 2001; Imbens and Lemieux, 2007; Lee and Lemieux, 2010) using the potential outcomes framework of Rubin (1974) and Holland (1986). In this framework, each establishment has two potential outcomes: the election passes, indicated by $W_i = 1$, and the election fails, indicated by $W_i = 0$. Occupational safety for each outcome is denoted $Y_i(W_i)$. The causal effect of unionization on occupational safety for each establishment is $Y_i(1) - Y_i(0)$, and the causal effect among all establishments is $E[Y_i(1) - Y_i(0)]$.

The empirical strategy identifies the causal effect of unionization specifically among establishments with a positive vote share of 50 percent, the cutoff for a successful union election. The vote share relative to the cutoff is denoted $X_i$. In this case, $W_i = 1$ for all establishments with $X_i > 0$, and $W_i = 0$ for all establishments with $X_i \leq 0$. The key identification assumption is that the conditional expectation functions $E[Y_i(1)|X_i]$ and $E[Y_i(0)|X_i]$ are smooth around the cutoff zero, requiring both observable and unobservable characteristics that affect $Y$ to be smooth at the cutoff. If so, the estimand for the causal effect among establishments at the cutoff, $E[Y_i(1) - Y_i(0)|X_i = 0]$, is given by $\lim_{x \downarrow 0} E[Y_i(1)|X_i] - \lim_{x \uparrow 0} E[Y_i(0)|X_i]$.

The causal effect is estimated using the following regression discontinuity model:

$$Y_i = \alpha + \beta W_i + G_F(X_i) + W_i G_P(X_i) + \epsilon.$$  \hspace{1cm} (1)
The terms $G_F(X_i)$ and $G_P(X_i)$ are polynomial functions of $X_i$, excluding the intercepts. Thus, $E[Y_i(0)|X_i]$ is modeled as $\alpha + G_F(X_i)$ for $X_i \leq 0$, and $E[Y_i(1)|X_i]$ is modeled as $\alpha + \beta W_i + G_F(X_i) + W_i G_P(X_i)$ for $X_i > 0$. The coefficient of interest is $\beta$, which represents the discontinuity in $Y$ at the cutoff. If the identification assumptions are valid, $\beta$ measures the causal effect of unionization on occupational safety among establishments at the cutoff.

4 Data

Data on union elections come from the NLRB. The data were downloaded from the online repository www.data.gov, which contained data releases for years 1999 to 2010. The data are recorded at the establishment level and report the establishment name, address, and industry. Regarding the election, the data report the number of eligible voters, valid votes cast, and votes for and against unionization.

Data on occupational safety come from OSHA’s Data Initiative (ODI). The ODI began in 1996 as part of OSHA’s Site Specific Targeting (SST) plan, which was developed to better target more dangerous establishments for an OSHA inspection. Case rates were collected directly from employers through the ODI, and these case rates were used to target high case-rate establishments for inspection. From 1996 to 2011, the ODI surveyed approximately 80,000 establishments annually, with a goal of surveying all establishments meeting the survey criteria at least once every three years. The ODI largely excluded establishments in construction and establishments with fewer than 40 employees. The data are recorded at the establishment level and report the establishment name, address, and industry. Regarding occupational safety, the ODI reports two accident case rates. The first is the total case rate (TCR), which includes cases involving death, days away from work, job restrictions, job transfers, and medical attention beyond first aid. The second is a subset of the TCR.

\footnote{Using ODI data and exploiting the Site Specific Targeting Plan, Li and Singleton (2017) examine the effect of workplace inspections on worker safety.}

\footnote{In 1997, this threshold was increased to 60 employees.}
and includes only cases involving days away from work, job restrictions, and job transfers (DART). Both rates are reported annually per 100 full-time equivalent workers.

The pooled NLRB data contain 40,163 closed records of attempts at union representation. These records were matched to each calendar year of the ODI based on the establishment name and address. According to the matching procedure, 8,580 records from the NLRB had at least one observation in the ODI. A low match rate is expected, as the ODI is not representative of all establishments. The data are restricted to 5,911 records with a vote date, eliminating records with withdrawn petitions, and then to 5,714 records with votes in years 1999 to 2010. The data are further restricted to 4,130 records with at least 20 valid votes. This is standard in related studies, as a larger number of voters increases the uncertainty of the election outcome (DiNardo and Lee, 2004; Frandsen, 2014). Finally, records are omitted if the vote share is missing. The final sample contains 4,118 NLRB elections from 1999 to 2010 with 19,752 unique ODI observations from 1996 to 2011. Once stacked, the data are an unbalanced panel, since establishments are not observed in every year of the ODI.

One reason an establishment may not be observed in the ODI is that it exists, but was not surveyed in a particular year. Another reason is that an establishment did not yet exist several years before the election or dissolved sometime after the election. Firm dissolution is potentially problematic for the empirical strategy if affects establishment survival differentially by occupational safety. For example, if a union election causes more dangerous establishments to dissolve, then average safety would improve over time, independent of within-firm changes in safety. To examine firm survival, the NLRB records are matched to a national database of over 35 million establishments a year, compiled annually by InfoUSA.4

The empirical strategy implicitly assumes that all establishments with a successful union election ultimately unionize. To support this assumption, the election data are

---

4The Infogroup compiles this information by first detecting businesses through numerous sources, such as over 4,300 yellow and white pages, county-level public sources, utility connects and disconnects, real estate tax assessor data and web research. It then calls every company in the United States every year.
matched to “notices of bargaining” data from the Federal Mediation and Conciliation Service (FMCS). A notice is filed to create, terminate, or modify a labor contract. Thus, a notice is an indicator of union activity, but not necessarily of unionization. The FMCS data are available for years 1997 to 2016.

5 Sample Summary

Summary statistics of the analysis sample are reported in Table 1. The vote share is calculated as the share of valid votes in favor of unionization relative to the 50 percent cutoff. Industry is determined by the NAICS codes reported in the NLRB. 42.5 percent of establishments are in manufacturing, and 23.4 percent are in health services. An additional 5.03 percent are in construction, 8.60 percent are in transportation, and the remaining 20.45 percent are in other industries. The share of establishments in health services is greater among successful elections (34.7 percent relative to 14.9 percent), and the share of establishments in manufacturing is greater among failed elections (48.9 percent relative to 34.0 percent). The geographic region is determined by the state reported in the NLRB. The smallest share of elections was in the South (20.3 percent), and the greatest share was in the Midwest (32.9 percent). However, elections in the Midwest represented the greatest share of failed elections (36.3 percent). The final two rows report case rates from the ODI in the calendar year of the election, restricted to establishments that are matched to the ODI in that year. The TCR was 13.2, and the mean DART rate was 7.99. Both rates are higher among establishments with failed elections, and the differences are statistically significant at the 5 percent level.

The identification strategy assumes that the conditional expectation functions are smooth near the cutoff. Support for the assumption is provided in two ways. First, Figure 1 plots the distribution of the vote share. As shown, the distribution is slightly skewed to the

5DiNardo and Lee (2004) similarly match union election data to the FMCS data.
right, with a mass at 0.5. Importantly, there does not appear to be excessive bunching above or below the cutoff. Bunching would suggest “gaming” near the cutoff and could potentially violate the smoothness assumption of the conditional expectation functions.

Second, Figure 2 illustrates the share of establishments in manufacturing and health services by 20 evenly-spaced bins of the vote share. The figure also plots the relationship between of industry and vote share using local linear regression, estimated separately above and below the cutoff. As shown, industry composition appears smooth near the cutoff, consistent with the identification assumption.

The discontinuity in industry can be estimated using equation (1), with indicators of industry as outcome variables. The discontinuity estimates are reported in the first two rows of Table 2. Each row corresponds to a different outcome variable, and each row reports six different estimates using first and second order polynomials across three bandwidths. As shown, none of the discontinuity estimates are statistically significant, and most are less than two percentage points in absolute value. The exceptions are in the final two columns, where the bandwidth ranges from -0.2 to 0.2. The mean squared error is smallest in column three for manufacturing and column six for health services, where the estimated discontinuities are 1.9 percentage points and -3.1 percentage points, respectively.

6 Results

6.1 Union Activity

The identification strategy also assumes that establishments with a successful union election are more likely to unionize, particularly at the cutoff. The first panel in Figure 3 plots the share of establishments that match to a notice of bargaining in the FMCS. The match rates are calculated by calendar year before and after the union election, with period zero corresponding to the calendar year of the election. Among establishments in which the
election passed, the match rate increases sharply in the calendar year of the election and
the year after, consistent with an increase unionization. Among establishments in which the
election failed, there is no change in the match rate in the year of the election relative to the
pre-existing trend.

To examine union activity at the cutoff, Figure 4 illustrates match rates by vote
share. There is no visible discontinuity in the match rate at the cutoff before the election, as
shown in the first panel. However, there is a discontinuous increase in the match rate at the
cutoff in the calendar year of the election and the year after, as shown in the second panel.
Estimates of the discontinuity using equation (1) are reported in the third and fourth rows
of Table 2. Before the election, there is no discontinuity in the match rate at the cutoff.
After the election, the discontinuity estimates range from 18.6 to 39.2 percentage points, all
of which are statistically significant at the one percent level. The sensitivity of the estimates
reflect, in part, that the match rate increases, then plateaus, just above the cutoff. Taken
together, the estimates confirm that establishments just above the cutoff are more likely to
exhibit union activity.

6.2 Firm Survival

To examine establishment survival, the second panel of Figure 3 plots the share of
establishments that match to the InfoUSA data in each calendar year relative to the election.
In the year of the election, the match rate is 75.4 percent among establishments in which the
election passed and 71.8 percent among establishments in which the election failed. The
less-than-perfect match rate may be due to missing establishments in the InfoUSA database,
measurement error in the establishment name and address, or the matching procedure. From
period zero, the match rates then decline with each period before and after the election. By
period five, the match rates reach 42.4 percent and 44.3 percent, respectively. Because the

---

6In comparison, the match rate of NLRB data records to the Census Bureau’s Longitudinal Business Database was 82 percent in Frandsen (2014).
InfoUSA is intended to encompass all establishments, the decline in match rates most likely reflects that some establishments no longer exist.

To examine firm survival at the cutoff, Figure 5 illustrates match rates by vote share for each period from -4 to 4. As shown, there is no large discontinuity in the match rate in any period.\footnote{Using election data matched to the Census Bureau's Longitudinal Business Database, Frandsen (2014) finds a negative discontinuity in establishments survival following union elections.} Estimates of the discontinuity in periods zero and four using equation (1) are presented in Table 2. As shown, none of the discontinuity estimates are statistically significant, though some estimates are sizeable depending on the model specification. Given these estimates and the graphical evidence in Figure 5, there does not appear to be a discontinuity in firm survival at the cutoff.

Another indication of firm survival is the match rate to the ODI. Although the match rate will generally be low in a single period, since the ODI did not survey all establishments meeting the survey criteria every year, a decline in the match rate over time would be indicative firm dissolution. The third panel of Figure 3 plots the share of establishments that match to the ODI in each calendar year relative to the election. As shown, the match rate is greatest in the year of the election, when the establishment is known to exist, and then declines with each period before and after the election.

Figure 6 illustrates the mean match rate by vote share during periods before and after the election. Periods -3, -2, and -1 are used for the pre-election period, and periods 1, 2, and 3 are used for the post-election period. The data are pooled here because they similarly pooled to examine accident case rates below. As shown, there is no large discontinuity in the mean match rate at the cutoff before or after the election. Discontinuity estimates are presented in the last rows of Table 2. None of the estimates are statistically significant, though they are imprecise and sensitive to the model specification. Again, given these estimates and the graphical evidence in Figure 6, there does not appear to be a discontinuity in firm survival at the cutoff.

It is important to note that matching to the InfoUSA is not systematically related
to matching to the ODI. For example, the ODI match rate in the year of the election is 38.8 percent among establishments that match to InfoUSA and 36.2 percent among establishments that do not. This suggests that the non-matches are likely due to measurement error and the matching process, rather than the non-existence of an establishment. Thus, when examining case rates, all matches to the ODI are utilized, not just matches that are contemporaneously matched to the InfoUSA.

6.3 Case Rates

The effect of unionization on occupational safety is examined using data from the ODI. To examine occupational safety before and after the election, ODI observations are pooled across multiple periods. Periods -3, -2, and -1 are used for the pre-election period, and periods 1, 2, and 3 are used for the post-election period. The data are pooled to increase the number of observations, since the ODI did not survey all establishments meeting the survey criteria every year. Moreover, the analysis period is limited to just three calendar years before and after the election to limit the potential bias due to differential firm survival at the cutoff.

Case rates by vote share are illustrated in Figures 7 and 8. Figure 7 illustrates the TCR, and Figure 8 illustrates the DART rate. In each figure, the first panel corresponds to before the election, and the second panel corresponds to after the election. In all four panels, the mean case rate appears to decrease below a vote share of -0.3, then trend monotonically thereafter. These trends appears to increase slightly in slope from before to after a union election, suggesting a relative increase in reported case rates among establishments with greater vote shares. However, there does not appear to be discontinuity in case rates at the discontinuity for either case rate before or after the election.

Discontinuity estimates from equation (1) are reported in Tables 3 and 4. Table 3 provides estimates before the election, and Table 4 presents estimates after the election. The tables present estimates for both the TCR and DART rate using different bandwidths
and polynomials and with and without covariates. The covariates consist of fixed effects for
calendar year, industry, and state.

As shown, none of the estimates are statistically significant. Moreover, all estimates
are less than one in absolute value, with some exceptions before the election. The estimates
are more precise for the DART than for the TCR. This may reflect that cases involving
medical attention beyond first aid - which is included in the TCR but not the DART - are
more subjective than cases involving days away from work. Regarding the six estimates for
the DART after election with covariates, five can reject an effect of -1.5, and one can reject
an effect of -1.0. Conversely, four can reject an effect of 2.0, and three can reject an effect
of 1.0. Thus, the effect of unionization on the DART rate appears limited, with estimates
centered around zero.

The estimates are more precise when comparing all establishment just above and
below the cutoff, not just at the discontinuity. For example, by comparing establishments
within ten percent above the cutoff to establishments within ten percent below, the estimated
difference in the DART is 0.177, with a confidence interval of -0.504 to 0.857, and the
estiamted difference in the TCR is 0.461, with a confidence interval of -0.24 to 1.17.\(^8\) For
these estimates to represent causal effects, a sufficient assumption is that the conditional
mean functions are constant within the bandwidth of analysis. This assumption is supported
by the near-zero slope of the linear regression estimates illustrated near the cutoff in Figures 7
and 8.

7 Conclusion

This study attempts to identify the effects of unionization on occupational safety.
For identification, the empiricial strategy focuses on narrow union election for identification
using a regression discontinuity model. Data on occupational safety come from the Oc-

\(^8\)These are regression-adjusted differences, controllling for calendar year, industry, and state.
cupational Safety and Health Administration, which contain accident case rates reported by employers at the establishment level. According to the results, union activity increased sharply at the cutoff for a successful union election, but there was relatively no change in accident case rates.

This study makes two main contributions to the related literature. First, this study is the first to identify the effect of labor unions on occupational safety by exploiting the timing of union elections. By following the same establishments over time, the empirical strategy addresses the issues of selection, whereby more dangerous firms are more likely to unionize. Second, this study is one of only a few studies that provide suggestive evidence that labor unions improve occupational safety. Indeed, most studies in the literature conclude that unionization is associated with greater accidents and injuries (Donado, 2015). This study also has implications for the literature on unions and wages, since improvements in occupational safety may decrease wages through compensating differentials. If so, the union wage premium may understate the total welfare gains from collective bargaining.

A limitation of this study is that, because case rates are self-reported, the mechanism for the results is unclear. The issue is that unionization may pressure employers to accurately record and report accident case rates, thereby increasing case rates. Thus, the results may reflect that unionization has no effect on occupational safety or reporting, or that unionization simultaneously improves occupational safety and increases reporting. Thus, it is important to consider more objective measures of occupational safety in future research.
Appendix

The data used in this study are compiled from several sources matched at the establishment level. The data on union election come from the NLRB; the data on firm survival come from the InfoUSA; and the data on union activity come from the FMCS. This section describes the procedures to match the data across sources.

The NLRB data contain union elections held from 1999 to 2010. The analysis sample includes closed cases of representation petitions in manufacturing with 20 or more valid votes. To link the NLRB data to other datasets, the establishment name and address are standardized. For the establishment name, all the special characters and common words, such as company, limited, and corporation, are first removed. If the listed formal name is different from the case name, or the establishment is “doing business as (DBA) one under another name, the second name is saved in a separate variable, and both names are used when linking the election records to other datasets. In the street address variable, all the special characters, and the floor, suite, and room numbers are removed. The common words, such as street, avenue, and road, are replaced with the abbreviations. The city names in the election records are compared to a list of all the city names in the US from Census, and the unmatched city names are manually checked for any misspelling.

To examine the effect of union on workplace safety, the NLRB data are linked to establishment level injury rates from ODI and to records on notices of bargaining from FMCS. The establishment name and address in the ODI and FMCS are first standardized using the same method when standardizing NLRB data. The information used to match across datasets include establishment name, street, address, city, state, and zip code. The matching is conducted in several steps using criteria from the exact match to more relaxed criteria. The first round of matching is based on exact match of name, street, city, and state. If a record in ODI or FMCS is matched to a union election, it is removed and not included in further rounds of matching. The second round of matching is based establishment name, zip code, city, and state. Then the records are matched based on the first six letters of the
name and address. The process is repeated for establishments with multiple names.

To construct the indicator of firm survival, the NLRB data is matched to InfoUSA. InfoUSA provides an annual list of all the active establishments operated in US from 1997 to 2013. Each establishment has a unique identifier, which can be used to link establishments across years. The election records are first matched to the records in InfoUSA from the calendar year of election using the matching procedures described above. If an establishment is matched to a record in InfoUSA in the same calendar year, the unique identifier of the establishment is used to check if the establishment has a record in InfoUSA in years prior to and after the election. The establishments in InfoUSA report the year of establishment, which is also used to determine the existence of an establishment in years prior to the election.
References


<table>
<thead>
<tr>
<th>Election Outcome</th>
<th>All</th>
<th>Pass</th>
<th>Fail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vote Share</td>
<td>-0.0002</td>
<td>0.215</td>
<td>-0.162</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>0.425</td>
<td>0.340</td>
<td>0.489</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.011)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Health Services</td>
<td>0.234</td>
<td>0.347</td>
<td>0.149</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.011)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Northeast</td>
<td>0.295</td>
<td>0.332</td>
<td>0.266</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.011)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Midwest</td>
<td>0.329</td>
<td>0.283</td>
<td>0.363</td>
</tr>
<tr>
<td></td>
<td>(0.007)</td>
<td>(0.011)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>South</td>
<td>0.203</td>
<td>0.192</td>
<td>0.212</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.009)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Total Case Rate</td>
<td>13.2</td>
<td>12.5</td>
<td>13.7</td>
</tr>
<tr>
<td></td>
<td>(0.276)</td>
<td>(0.393)</td>
<td>(0.382)</td>
</tr>
<tr>
<td>DART Rate</td>
<td>7.99</td>
<td>7.59</td>
<td>8.29</td>
</tr>
<tr>
<td></td>
<td>(0.182)</td>
<td>(0.251)</td>
<td>(0.255)</td>
</tr>
<tr>
<td>Observations</td>
<td>4,118</td>
<td>1,769</td>
<td>2,349</td>
</tr>
</tbody>
</table>

The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. The vote share is calculated as the percent of eligible voters in favor of unionization relative to the 50 percent cutoff. The total case rate includes cases involving death, days away from work, job restrictions and transfers, and medical attention beyond first aid. The DART rate includes cases involving days away from work and job restrictions and transfers. The rates are calculated during the year of the election, using only establishments that match to the ODI that year, and are reported per 100 full-time equivalent workers. Standard errors are in parentheses.
<table>
<thead>
<tr>
<th>Outcome Polynomials</th>
<th>Bandwidth [0.1, 0.1]</th>
<th>Bandwidth [0.15, 0.15]</th>
<th>Bandwidth [0.2, 0.2]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>0</td>
<td>0.015</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.083)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Health Services</td>
<td>0</td>
<td>0.013</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>(0.046)</td>
<td>(0.070)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>FMCS</td>
<td>-2, -1</td>
<td>0.041</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.031)</td>
<td>(0.047)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>FMCS</td>
<td>0, 1</td>
<td>0.277**</td>
<td>0.186**</td>
</tr>
<tr>
<td></td>
<td>(0.045)</td>
<td>(0.069)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>InfoUSA</td>
<td>-4</td>
<td>-0.061</td>
<td>-0.034</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.081)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>InfoUSA</td>
<td>0</td>
<td>-0.071</td>
<td>-0.062</td>
</tr>
<tr>
<td></td>
<td>(0.048)</td>
<td>(0.074)</td>
<td>(0.040)</td>
</tr>
<tr>
<td>InfoUSA</td>
<td>4</td>
<td>-0.102</td>
<td>-0.05</td>
</tr>
<tr>
<td></td>
<td>(0.055)</td>
<td>(0.085)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>ODI</td>
<td>-3, -2, -1</td>
<td>0.020</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.062)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>ODI</td>
<td>1, 2, 3</td>
<td>0.034</td>
<td>0.049</td>
</tr>
<tr>
<td></td>
<td>(0.040)</td>
<td>(0.062)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,374</td>
<td>1,374</td>
<td>1,989</td>
</tr>
</tbody>
</table>

The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. Each figure is a discontinuity estimate derived from a regression discontinuity model. The bandwidth is measured by the vote share is calculated as the percent of eligible voters in favor of unionization relative to the 50 percent cutoff. FMCS indicates a match to the Federal Mediation and Conciliation Service; InfoUSA indicates a match to the InfoUSA database; ODI indicates a match
to OSHA’s Data Initiative. Periods are defined relative to the calendar year of the election. Standard errors are in parentheses. * and ** indicate significance at the five and one percent level, respectively.
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Periods</th>
<th>Controls</th>
<th>Bandwidth [-0.1,0.1]</th>
<th>Polynomial 1</th>
<th>Polynomial 2</th>
<th>[-0.15,0.15]</th>
<th>Polynomial 1</th>
<th>Polynomial 2</th>
<th>[-0.2,0.2]</th>
<th>Polynomial 1</th>
<th>Polynomial 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>TCR</td>
<td>-3, -2, -1</td>
<td>No</td>
<td></td>
<td>0.021</td>
<td>-0.580</td>
<td>-1.071</td>
<td>-0.474</td>
<td>-0.812</td>
<td>-0.923</td>
<td>(1.216)</td>
<td>(1.903)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.216)</td>
<td>(1.903)</td>
<td>(1.016)</td>
<td>(1.533)</td>
<td>(0.883)</td>
<td>(1.311)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCR</td>
<td>-3, -2, -1</td>
<td>Yes</td>
<td></td>
<td>0.635</td>
<td>0.451</td>
<td>0.0670</td>
<td>0.383</td>
<td>0.187</td>
<td>0.343</td>
<td>(1.166)</td>
<td>(1.819)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.166)</td>
<td>(1.819)</td>
<td>(0.971)</td>
<td>(1.475)</td>
<td>(0.850)</td>
<td>(1.256)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DART</td>
<td>-3, -2, -1</td>
<td>No</td>
<td></td>
<td>0.997</td>
<td>1.633</td>
<td>0.0621</td>
<td>1.423</td>
<td>0.403</td>
<td>0.458</td>
<td>(0.712)</td>
<td>(1.115)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.712)</td>
<td>(1.115)</td>
<td>(0.636)</td>
<td>(0.959)</td>
<td>(0.565)</td>
<td>(0.839)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DART</td>
<td>-3, -2, -1</td>
<td>Yes</td>
<td></td>
<td>1.216</td>
<td>2.042</td>
<td>0.630</td>
<td>1.796</td>
<td>0.723</td>
<td>1.142</td>
<td>(0.699)</td>
<td>(1.092)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.699)</td>
<td>(1.092)</td>
<td>(0.630)</td>
<td>(0.956)</td>
<td>(0.558)</td>
<td>(0.825)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td></td>
<td></td>
<td></td>
<td>1,391</td>
<td>1,391</td>
<td>2,015</td>
<td>2,015</td>
<td>2,595</td>
<td>2,595</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. Each figure is a discontinuity estimate derived from a regression discontinuity model. The bandwidth is measured by the vote share is calculated as the percent of eligible voters in favor of unionization relative to the 50 percent cutoff. The total case rate includes cases involving death, days away from work, job restrictions and transfers, and medical attention beyond first aid. The DART rate includes cases involving days away from work and job restrictions and transfers. Periods are defined relative to the calendar year of the election. Standard errors are in parentheses. * and ** indicate significance at the five and one percent.
### Table 4: Discontinuity Estimates - Case Rates After Election

The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative.
The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. The figure plots the percent of establishments in an industry by bins of the vote share relative to the 50 percent cutoff. The line is derived from local linear regression, estimated separately above and below the cutoff.
Figure 3: Dynamic Match Rates by Election Outcome

The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. The figure plots the match rate of establishments to the FMCS, InfoUSA, and the ODI in each calendar year relative to the election. FMCS is the Federal Mediation and Conciliation Service; InfoUSA is the InfoUSA database; ODI indicates is OSHA’s Data Initiative.
The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. The figure plots the match rate of establishments to notices of bargaining from the Federal Mediation and Conciliation Service by bins of the vote share relative to the 50 percent cutoff. The line is derived from local linear regression, estimated separately above and below the cutoff.
The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. The figure plots the match rate of establishments to the InfoUSA by bins of the vote share relative to the 50 percent cutoff. The line is derived from local linear regression, estimated separately above and below the cutoff.
The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. The figure plots the match rate of establishments to data from OSHA’s Data Initiative by bins of the vote share relative to the 50 percent cutoff. The line is derived from local linear regression, estimated separately above and below the cutoff.
The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. The figure plots the mean TCR by bins of the vote share relative to the 50 percent cutoff. The total case rate includes cases involving death, days away from work, job restrictions and transfers, and medical attention beyond first aid. The line is derived from local linear regression, estimated separately above and below the cutoff.
The sample is derived by matching data from union elections in years 1999 to 2010 to data from OSHA’s Data Initiative. The figure plots the mean DART rate by bins of the vote share relative to the 50 percent cutoff. The DART rate includes cases involving days away from work and job restrictions and transfers. The line is derived from local linear regression, estimated separately above and below the cutoff.