Unionization, Employer Opposition, and Establishment Closure

Sean Wang†  Samuel Young†
(Job Market Paper)

November 22, 2021

Revised Often: Latest Version Here

Abstract

We study the effect of private-sector unionization on establishment employment and survival. Specifically, we analyze National Labor Relations Board (NLRB) union elections from 1981 to 2005 using administrative Census data on the universe of establishments in the U.S. Our research design combines difference-in-differences and regression discontinuity extrapolation methods to estimate treatment effects that include elections that win by larger margins of support. We show that unionization decreases an establishment’s employment and likelihood of survival. We hypothesize that two reasons for these effects are firms’ ability to avoid working with new unions and managers’ opposition to unions. We test this hypothesis for unionization in manufacturing, the largest sector where we find substantial negative effects. There, the negative effects are significantly larger for elections at multi-establishment firms, especially those with no other unionized establishments. We provide direct evidence suggesting that some of these differences are driven by multi-establishment firms shifting employment from newly unionized establishments to other establishments. Finally, we use the length of delays during the election process as a proxy for managers’ opposition to the union and find substantially larger effects of elections with longer delays. Taken together, our results are consistent with firms’ union avoidance tactics playing a role in explaining the overall negative effects of unionization.

*We are grateful to Daron Acemoglu, David Autor, and Simon Jäger for guidance and advice throughout this project. We thank Josh Angrist, Jon Cohen, David Hughes, Sylvia Klosin, Tom Kochan, Felix Koenig, Mike Piore, Frank Schilbach, Garima Sharma, Martina Uccioli, John Van Reenen, Michael Wong, and Josef Zweimüller and seminar participants at the U.S. Census Bureau for helpful comments. This paper benefited greatly from Henry Hyatt and Kirk White’s data expertise. We thank Stephanie Bailey, Jim Davis, and Nathan Ramsey for assistance with the data access and the disclosure process. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. (1745302). All errors are our own. Disclaimer: Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau’s Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2389. (CBDRB-FY22-P2389-R9311 and CBDRB-FY22-P2389-R9358)

†Massachusetts Institute of Technology. Emails: swang1@mit.edu and sgyoung@mit.edu.
1 Introduction

Union elections in the U.S. are extremely contentious. Managers frequently threaten to close establishments if they unionize, and survey evidence suggests that some follow through on these threats (Bronfenbrenner, 1996). The conventional economic explanation for establishments closing after unionization is that unions’ demands for wage increases or workplace changes may make businesses unprofitable. This explanation, however, is not supported by existing research, which finds little evidence of successful union elections leading to higher wages or lower productivity.

Consider two examples of how employers respond to unions that suggest alternative reasons why unionization may lead to establishment closure. First, during a 2017 campaign to unionize the news website Gothamist, the owner wrote to employees, “as long as it’s my money that’s paying for everything, I intend to be the one making the decisions.” One week after the workers voted to unionize, the owner shut down the business (Wamsley, 2017). This example suggests that some establishment closures could be driven by managers’ unwillingness to operate alongside unions or their general dislike of working with unions. Second, consider Boeing’s production of 787 airplanes. In 2011, Boeing shifted production of some 787s away from a unionized plant in Washington state to a non-union plant in South Carolina. Today, it manufactures all 787s in South Carolina (Cameron, 2020). According to a Boeing executive, the motivation “was not the wages we’re paying today. It was that we cannot afford to have a work stoppage, you know, every three years” (Greenhouse, 2011). This example suggests that conflictual labor relations between unions and firms may lead to establishment closures. Additionally, the example illustrates how some firms can avoid working with unions while continuing to operate.

This paper assesses whether these examples generalize by analyzing the effect of unionization on establishment employment and survival in the U.S. We then examine whether firms’ ability to avoid working with new unions and managers’ general opposition to unions help explain the effects of unionization. Our setting is around 27,000 private-sector union certification elections through the National Labor Relations Board (NLRB) from 1981-2005. We link these elections to administrative Census Bureau data on establishment employment, survival, and total factor productivity and union contract data from the Federal Mediation and Conciliation Service (FMCS).

We analyze these elections using a novel research design that extends standard difference-in-differences techniques with falsification tests from the regression discontinuity extrapolation literature. This strategy allows us to estimate treatment effects that include elections that win by large margins of support. We find that unionization decreases establishments’ employment primarily...
by lowering their likelihood of survival. Five years after an election, we estimate a survival effect of four percentage points (pct. pts.) relative to a survival rate of 82% for establishments where the union lost. The employment effects are also bigger for larger margin-of-victory elections.

Motivated by these overall effects of unionization, we test whether firms’ union avoidance tactics and managers’ opposition to unions help explain the effects. For this analysis, we focus on manufacturing elections for two reasons. First, we have better data to test the specific parts of our hypothesis.\footnote{The manufacturing data are better for two reasons. First, in manufacturing, a prominent union avoidance method is shifting employment away from newly unionized establishments (Bluestone and Harrison, 1982; Verma, 1985; Kochan et al., 1986a) which we can measure with the establishment and firm linkages. In other industries, such shifting may be more difficult, and other union avoidance tactics are harder to measure with our data (e.g., replacing unionized workers with independent contractors or opening a separate non-union firm (Evans and Lewis, 1989; Hatton, 2014)). Second, we have high-quality plant-level productivity measures in manufacturing that we use to test effect heterogeneity by baseline productivity.} Second, we find that manufacturing is the largest sector with substantial negative effects. To show this, we estimate separate effects for three broad industry groups: manufacturing, services, and other industries.\footnote{Our services elections are primarily in health, retail trade, admin and support, accommodation and food, and information and the “other” industries are primarily transportation, wholesale trade, construction, and warehousing.} We find that the overall employment and survival declines are driven by similarly sized effects in manufacturing and “other industries”. The ten-year effect on survival for manufacturing elections is eight pct. pts.. Conversely, the estimates are much smaller and sometimes not significantly different than zero for unionization in the service sector.

The first part of our hypothesis is that firms avoid working with unions by shifting production away from newly unionized plants to other plants. To test this, we start by estimating whether the effects of union elections are larger at establishments part of multi-establishment (or multi-unit, MU) firms (e.g., an election at one of the plants for a firm with several plants) versus single-establishment (SU) firms. Consistent with the hypothesis, we find significantly larger employment and survival effects for elections at MU firms than SU firms. The ten-year effects on survival are twelve pct. pts. versus three pct. pts. for MU and SU firms, respectively.

Next, we directly test whether MU firms shift production away from newly unionized plants to other plants. For all elections at MU firms, we compare the employment and survival of the firms’ other manufacturing plants after a successful versus unsuccessful union election. Overall, we do not find differences in post-election outcomes between the firms’ other plants. Yet, this is not surprising since these plants may produce very different products. To test for employment shifting across plants that produce more similar products, we focus on the firms’ plants in the same three-digit NAICS industry as the election plant. With this restriction, we find significantly higher employment growth for plants at firms with a winning election than plants at firms with a losing election. However, these short-run effects become insignificant five years after the election, at which point the firm may have shifted production to new non-unionized plants. Overall, this evidence supports firms shifting production away from newly unionized plants to avoiding working with new unions as one explanation for the negative impact of unionization.

The second part of our hypothesis is that the effects of unionization are greater when management is more opposed to the union. To test this, we estimate treatment effect heterogeneity
based on two proxies for managers’ opposition. First, we estimate effects separately for MU firms with and without any other unionized establishments. During this time, less unionized firms would more “vigorously resist dealing with unions” and that their managers’ anti-union philosophies were at times a key motivation for this resistance (Freedman, 1979; Foulkes, 1980). Additionally, similar to Selten (1978)’s “chain store paradox,” non-unionized firms may have a strong incentive to aggressively deter the first unionization attempt (even if not economically profitable when considering each establishment in isolation). Supporting these accounts, we find that the long-run negative employment and survival effects of unionization are significantly larger for elections at non-unionized firms than at (partially) unionized firms.

Our second proxy for managers’ opposition to unions is delays during the election process. Strategies that delay elections are a key way that managers attempt to influence elections and consequently a proxy for managers’ opposition. For example, in “Confessions of a Union Buster,” Levitt and Conrow (1993) write that the NLRA “presents endless possibilities for delays, roadblocks, and maneuvers that can undermine a union’s efforts” and that delay “steals momentum from a union-organizing drive.” We define election delay as the time between the date the union filed for an election and the election date and estimate separate treatment effects for elections with shorter and longer delays. We find significantly larger negative effects on employment and survival for elections with longer delays. For example, the three- and ten-year survival effects for MU elections in the top tercile of election delay times are 13 and 20 pct. pts. versus 5 and 7 pct. pts. for the bottom tercile.

Finally, we test whether the effects of unionization differ based on establishments’ baseline total-factor productivity (TFP), which we would expect under the conventional explanations for why unionization leads to establishment closures. Specifically, theoretical and empirical work from other contexts points to productivity as a key determinant of plant exit decisions. This suggests that wage increases or productivity declines from unionization should lead to larger survival effects for low productivity establishments.6 We do not, however, find significant differences in the employment or survival effects of unionization between establishments with different baseline TFP (we measure TFP using plant-level input and output data from the Annual Survey of Manufacturers). Thus, this evidence is more consistent with alternative explanations for why unionization leads to plant closures (e.g., our union avoidance hypothesis) than conventional explanations.

There are several potential interpretations of this evidence on managers’ opposition to unions and their use of union avoidance tactics. One interpretation is that the opposition and avoidance are motivated by a dislike of unions unrelated to the direct costs of unions. This interpretation helps resolve two puzzles alluded to earlier. First, existing research has not found that unionization raises wages or negatively affects establishments’ productivity. Second, we would not expect unions to make demands that push firms out of business, as this would directly harm their workers (Friedman, 1951). Under the conventional explanation for closures, both points are difficult to reconcile with our estimates of large survival effects. Yet, they are consistent with closures being driven by managers’

---

6See, for example, the theoretical and empirical literature on the minimum wage (Berger et al., 2021; Dustmann et al., 2020) and the literature on productivity being a key driver of plant exit (Foster et al., 2008).
idiosyncratic dislike of unions. Additionally, this interpretation is consistent with our finding of no heterogeneity by baseline establishment productivity. On the other hand, we cannot rule out that our proxies for manager opposition simply measure where unions would have been the costliest. Supporting this interpretation, there is research suggesting direct costs of unionization, including equity declines following successful union elections, and that unions generally reduce firm profits (see Lee and Mas (2012) and Freeman and Medoff (1984), respectively). Overall, while our results do not speak directly to the magnitude of the direct costs of unionization, they suggest that the employment and survival effects of unionization may be excessive relative to these costs.

We implement several falsification tests of our identifying assumption that support our results not being driven by unobserved shocks correlated with election outcomes. Our empirical strategy is a difference-in-differences design that compares outcomes before and after union elections at establishments where the union won versus lost. Our identifying assumption is that outcomes at establishments with different election vote shares but the same baseline characteristics would have followed parallel trends had no election occurred. To support this assumption, we first show that only conditioning on baseline employment and industry yields similar pre-election employment and payroll growth rates between establishments with winning and losing elections. This similarity holds when we add much richer baseline covariates and for up to ten years before the union elections.

To further support our design, we assess several testable implications of our identifying assumption that are possible since we observe the “running variable” (vote shares). These falsification checks extend tests from the regression discontinuity extrapolation literature to a panel-data setting (Angrist and Rokkanen, 2015; Bennett, 2020). First, we show that the similarity in pre-election employment growth rates holds between finer vote share groups (i.e., comparing 60 versus 70 % elections). Second, we show that establishments’ post-election employment growth rates and survival were similar between losing elections with different vote shares. If our treatment effects were driven by contemporaneous shocks correlated with vote shares, we would expect different post-election outcomes between losing elections with different vote shares. Overall, these tests support our design by showing that our identifying assumption holds for several sets of observations where we see untreated potential outcomes.

Our empirical strategy combines features of regression discontinuity (RD) and panel data methods that have previously been used to analyze union elections. Although RD methods are appealing due to their internal validity, there are some disadvantages in this setting. First, the effects of close union elections may be quite different than elections with larger margins of support (e.g., higher vote shares may lead to stronger “mandates to bargain”). Additionally, there is manipulation around the 50 % threshold that leads to pre-election discontinuities in establishment characteristics (Frandsen, 2017). Our paper uses the panel dimension of the data to account for selection into winning versus losing elections while expanding the bandwidth around the threshold to include all 20-80 % vote-share elections. The wider bandwidth addresses the above issues and gives us more

---

7See DiNardo and Lee (2004); Sojourner et al. (2015); Knepper (2020); Bradley et al. (2017) for RD analyses and Freeman and Kleiner (1990a); LaLonde et al. (1996); Lee and Mas (2012); Dube et al. (2016); Goncalves (2021) for panel data analyses.
power to estimate heterogeneous treatment effects. Relative to the other panel-data analyses, we better exploit observing the vote shares to implement tests of our identifying assumption.\textsuperscript{8}

Compared to the previous literature, our overall employment and survival estimates match Frandsen (2021)’s findings of short-run decreases in establishment employment for close union elections and complement his suggestive evidence of negative survival effects.\textsuperscript{9} Our estimates also confirm other research finding employment declines following successful union elections in specific sectors (Sojourner et al., 2015; LaLonde et al., 1996).\textsuperscript{10} As a consequence, our results contrast with the null effects of close union elections in DiNardo and Lee (2004) and other research finding no relationship between unions and business survival (Freeman and Kleiner, 1999). These differences are potentially due to our use of higher-quality establishment survival data. Finally, our evidence of larger employment declines for successful elections with higher vote shares complements Lee and Mas (2012)’s finding of larger stock market declines for higher vote share elections.

Our evidence supporting the manager opposition and union avoidance hypothesis is novel relative to the economics literature on union elections but supports other research about firms’ responses to elections. For example, Bronfenbrenner (2000, 2001) report similar results from a survey of elections in the 1990s. She finds survival effects of twelve pctl. pts. following successful elections. She also finds that plant closing threats were more common at the types of elections where we find larger survival effects (e.g., in manufacturing and at MU firms). In addition, our finding that manufacturing firms with the strongest anti-union campaigns were more likely to shut down plants after successful elections adds to the literature on firms’ overall union avoidance tactics (Freeman and Kleiner, 1990a; Kleiner, 2001; Flanagan, 2007). In particular, this result complements Ferguson (2008)’s finding that successful elections with unfair labor practice charges (another proxy for employers’ opposition) are less likely to reach a first contract. Finally, our output shifting evidence is consistent with firms becoming less unionized during this time through investing in and opening non-union plants (Bluestone and Harrison, 1982; Kochan et al., 1986a; Verma, 1985).

The rest of the paper is structured as follows. Section 2 describes the institutional details of collective bargaining and NLRB union elections in the U.S. Section 3 describes the election and Census data we analyze. Section 4 discusses our empirical strategy and tests of our identifying assumption. Section 5 presents the results from estimating our empirical strategy for all union elections and presents our industry-heterogeneity estimates. Section 6 provides multiple tests of our manager opposition and union avoidance hypothesis in manufacturing. Finally, Sections 7 and 8 discuss our results in the context of the previous literature on unionization and broader implications.

\textsuperscript{8}Lee and Mas (2012) and Frandsen (2021) present pre-trends and post-election outcomes across the vote-share distribution but do not use these estimates as formal tests of their identifying assumptions.

\textsuperscript{9}His survival estimates are differences in survival probabilities around the 50 % threshold and he writes that “a causal interpretation of the differences in survival probability should be made with caution.” Alternatively, we assume conditional parallel trends between a wider range of winning and losing elections rather than assuming that unconditional survival probabilities are unaffected by the sorting around the 50 % threshold.

\textsuperscript{10}LaLonde et al. (1996) analyze the employment and output effects of manufacturing union elections from 1977-1989 using a difference-in-differences design. Yet, their analysis differs from ours on several dimensions. First, they do not analyze establishment exit, which makes interpreting the results conditional on survival difficult. Additionally, due to their smaller sample size, their pre-trend estimates are often imprecise, making it difficult to assess the parallel trends assumption.
2 Unionization through NLRB Elections

Collective Bargaining in the United States

The National Labor Relations Act (NLRA) guarantees most workers in the U.S. the right to collective bargaining and action. Under the NLRA, when a union represents a group of workers, their employer is required to bargain with the union over the conditions of employment.\textsuperscript{11} This bargaining generally occurs at the establishment level. During negotiations, the union may go on strike or the employer may “lockout” workers to pressure the other party. The NLRA also created the National Labor Relations Board (NLRB), a quasi-judicial agency that administers union elections and enforces unfair labor practice violations. Much of the current U.S. policy discussion around organized labor today focuses on increasing representation at non-unionized establishments.\textsuperscript{12} Our results speak directly to the consequences of increasing such successful unionization.

The NLRB Union Election Process

The primary way for private-sector workers to gain union representation is through a secret-ballot NLRB election. The organizing drive is initiated by workers at the establishment, either on their own initiative or prompted by outreach from a union. The first step is getting cards indicating union support signed by workers in the proposed “bargaining unit” (i.e., the workers the union would represent). The bargaining unit generally only contains workers at a single establishment and can range from workers in a single occupation (e.g., delivery truck drivers) to all non-managerial workers. After gathering signatures from at least 30\% of the bargaining unit, the union files an election petition with the NLRB. The NLRB then confirms that the cards show sufficient support for the union, resolves any disagreements over the composition of the bargaining unit, and schedules the election. After the petition is filed, employers frequently attempt to delay the election to reduce union support (e.g., by contesting the workers in the bargaining unit) (Levitt and Conrow, 1993).

Before the election, the union and employer often actively campaign for and against union representation. Union organizers and pro-union workers can campaign by speaking with other workers and publicly showing solidarity among union supporters. Employers have many campaign tools at their disposal including “captive audience meetings” where employees are required to listen to anti-union messaging and one-on-one meetings with supervisors. Employers also frequently hire “union avoidance” consultants and law firms (Logan, 2002). Finally, although there are legal restrictions on firing pro-union workers and threatening to close unionized establishments, these tactics still occur (Weiler, 1983; Schmitt and Zipperer, 2009). If a majority of workers vote in favor of the union, the union is certified by the NLRB to represent the bargaining unit. After the union is certified, the employer is required to bargain “in good faith” with the union. But the parties

\textsuperscript{11}The goal of these negotiations is a contract. Contracts commonly specify wage and non-wage compensation for each job title, grievance procedures for disputes, policies for implementing layoffs (e.g., seniority rules), and promotion policies. Production choices (e.g., product pricing, technology investment, and plant closing) are typically left to management (Slichter et al., 1960).

\textsuperscript{12}For example, the currently debated Protecting the Right to Organize (PRO) Act of 2021 would limit employers’ ability to campaign against union elections and increase penalties for unfair labor practices during elections.
are not required to reach an agreement. If a contract is not reached one year after the union is certified, employees can vote out the union by holding a decertification election.

NLRB elections are the primary method through which private-sector workers in an establishment gain union representation. However, there are two reasons why some unionization occurs without an election. First, the NLRA does not cover all workers (General Accounting Office, 2002). Second, covered workers can gain union representation without an election through voluntary “card check” recognition. However, card check is much less common than elections.

Close Elections and Motivation Including Larger Margin of Support Elections

An advantage of our empirical strategy is that it does not rely on comparing only elections that barely won or lost. One motivation for this is that there is non-random sorting of elections just around the 50% threshold (i.e., “vote share manipulation”). Figure 2 Panel A plots the vote share distribution for the elections in our sample and shows manipulation around the 50% threshold previously documented by Frandsen (2017) (e.g., a missing mass of elections that barely win). Frandsen (2021) also documents that this manipulation leads to large differences in observable establishment characteristics across the threshold (e.g., 13-22% differences in employment).

Another motivation is that the treatment effect of unionization may depend on the election vote share. For example, Lee and Mas (2012) find that the negative stock market effects of unionization are larger for higher margin of victory elections. A first reason for this is that very close union elections are often followed by lengthy litigation before bargaining begins (e.g., debates over challenged votes that could change the election result). Figure 2 Panel B presents the average and median number of days between the election date and the case closing date (e.g., the date when the union is officially certified and represents the workers). The figure shows striking increases in this delay time for very close elections (e.g., the median (mean) time for elections that barely win is around 118 (223) days versus only 11 (57) days for elections that win with 60% of the votes). Since delays can dampen the unions’ bargaining power, this evidence suggests that the effects of very close elections may be different than higher vote share elections.

Second, for close union elections, firms may delay the bargaining process anticipating a future decertification election. Figure 2 Panel C provides evidence of this by showing the probability of each certification election experiencing a decertification election in the five years following the

---

13 In a review, CRS (2013) find that 56-85% of successful elections result in first contracts during the period we consider.

14 Some workers lack collective bargaining rights (e.g., independent contractors, some small business employees, domestic workers, and “agricultural laborers”). Other workers have collective bargaining rights but are not covered by the NLRA. For example, airline and railroad employees’ collective bargaining rights are covered by the Railway Labor Act. While agricultural laborers are excluded from the NLRA, some states provide these workers with bargaining rights through state statutes. Finally, public-sector workers’ bargaining rights are covered by various federal, state, and local statutes.

15 Schmitt and Zipperer (2009) estimate that from 1998-2003 60% of workers were organized through NLRB elections but assume that before then 90% of organizing occurred through elections.

16 It is difficult to detect manipulation from the vote-share density figure because of the discreteness of the running variable and since our sample includes elections with a small number of votes. Consequently, we plot elections with exactly 50% of votes separately to make the manipulation easier to see. Frandsen (2017) finds evidence of manipulation using formal tests that accommodate discrete running variables. Additionally, Figure A2 plots the density of vote shares for elections with more than 50 votes where it is clearer to see evidence of manipulation.
original election. It shows that more than twelve percent of very close elections experience a future decertification election compared to less than five percent of larger margin of victory elections suggesting that higher margin of victory elections may be more likely to reach first contracts. A final reason why the treatment effect of unionization may vary by vote share is that unions that win with more support may be able to more credibly threaten to strike. Figure 2 Panel D supports this by showing that within manufacturing, where strikes were generally more common, the probability of a works stoppage is increasing in the vote share for winning elections.\(^{17}\) Overall, these results show that several proxies for the unions’ bargaining power are increasing in the election vote share suggesting that the effects of unionization may also differ along this margin.

**Selection into Union Elections and the Determinants of Winning Elections**

Since our empirical design compares winning versus losing elections with similar baseline characteristics, it is helpful to review the literature on selection into holding and winning elections. This literature motivates which baseline characteristics we condition on and our additions tests of whether election vote shares are related to other unobservable shocks. For selection into elections, Dinlersoz et al. (2017) find that elections are more likely at larger, more productive, and younger establishments. We account for this selection by only comparing establishments that hold elections.

For election results, workers, employers, and other idiosyncratic factors could all influence whether the union wins. For our empirical strategy, a concern is that vote shares may be related to future establishment productivity shocks. On the one hand, workers may be more likely to vote for the union when they expect their establishment to become more profitable in the future leading to a positive bias between vote shares and establishment growth. Alternatively, firms that expect to become more productive may campaign harder against the union generating a negative bias.

Research on election outcomes finds that these factors all play some role. The most consistent finding is higher union win rates at smaller bargaining units (Heneman and Sandver, 1983; Farber, 2001). Win rates also vary substantially across industries. In the 2000s, the win rate in manufacturing was around 40% versus 60% for services. These factors motivate our baseline specification that just conditions on baseline industry and establishment size. In terms of the influence of employer versus union campaigns, Bronfenbrenner (1991) finds that “union tactic variables explain more of the variance in election outcomes than any other group” including employer tactics or characteristics. Yet, others find that firms’ anti-union campaigns are associated with lower win rates (Freeman and Medoff, 1984, p.234-235). To address the concern that firms’ anti-union campaigns lead to a negative bias between vote shares and establishment growth, we implement multiple tests of whether election vote shares seem correlated with future productivity shocks.

\(^{17}\)The works’ stoppages are from the FMCS from 1984-2005. They include both strikes and employer-initiated lockouts.
3 Election Data, Census Data, and Outcomes of Interest

For our analysis, we combine union election and contract data with establishment- and firm-level data from the U.S. Census Bureau. These data are uniquely suited to study union elections in the U.S. First, the data contain the universe of establishments, the level where most elections are held. Analysis of more aggregated data would include establishments not directly affected by the elections and attenuate the direct effects of unionization. Second, the Census has constructed high-quality longitudinal establishment linkages that allow us to separate real establishment exit from spurious exit due to administrative or legal changes (Haltiwanger et al., 2013). This is important because survival is a key outcome of interest. Finally, the rich establishment covariates allow us to compare similar winning versus losing elections (e.g., same size, age, and industry).

NLRB Union Election Data

We combine data from multiple sources to construct a comprehensive dataset of union elections from 1962 to 2018. Specifically, we use the datasets assembled by Henry Farber, J.P. Ferguson, and Thomas Holmes and publicly available data from the NLRB. For duplicate observations across datasets, we pick one observation for each NLRB case number (see Appendix Section C for details). Appendix Figure A1 shows that this procedure yields a similar number of cases each year to the total cases reported in the NLRB’s annual reports. The election data contain information on election votes that we use to define treatment. Additionally, they contain employers’ names and addresses that we use to match elections to the Census data. Finally, the data include the election petition filing date, the actual election date, and the election closing date. We define our treatment time based on the filing date of each election because this is the earliest date we observe for each election. Additionally, we use these dates to define the time between filing the election petition and holding the election, a key proxy for managements’ opposition to the elections that we describe further in Section 6.

FMCS Contract Notice Data

To measure successful contract negotiations following elections and the presence of pre-existing unions, we use contract notice data from the Federal Mediation and Conciliation Service (FMCS) from 1984-2019. We combine data from Thomas Holmes and the FMCS. They include both notices of initial contracts (i.e., first-contract negotiation after an election) and contract renegotiation or reopening for existing contracts. These “notices of bargaining” are provided to the FMCS so it can be ready to provide mediation. Although filing is legally incentivized, underreporting is possible. These data also include names and addresses for matching. We use these data to measure whether an election establishment has any other workers covered by a collective bargaining agreement and whether the election establishment’s firm has any other unionized establishments.
Employment, Payroll, and Survival Data from the LBD

Our primary source of establishment-level outcomes is the Longitudinal Business Database (LBD). It contains annual employment and payroll data for the universe of non-farm, private sector establishments from 1976-2015 (Jarmin and Miranda, 2002). Our employment measure is the total number of employees who received wages or other compensation during the pay period that included March 12th. The payroll measure is the total “wages, tips, and other compensation” for employees over the entire year. Additionally, the data contain high-quality longitudinal establishment linkages that allow us to follow establishments over time, including across changes in firm ownership, and precisely define establishment survival. We define establishment survival based on the last year the establishment has non-zero employment. Finally, we use the Fort and Klimek (2016) 2012 North American Industrial Classification (NAICS) codes to classify each establishment into time-consistent industries across the entire time period.

Some complications with assigning employment and payroll to establishments part of multi-establishment (MU) firms are relevant for our analysis. In particular, although the LBD is at the establishment level, some of the annual employment and payroll data are received at higher levels of aggregation. These data need to be allocated across establishments (see Appendix Section C for details). The Census uses other establishment-level surveys to improve the establishment-level measures but these surveys are not conducted annually for all establishments so there may be some lags in these improvements. Consequently, if a unionized establishment at a MU firm shrinks or exits, some of this decrease in employment may be initially allocated to the firm’s other establishments creating a short-run underestimate of the effect of unionization at MU establishments. These allocation issues, however, are less of a concern for our longer-run outcomes (e.g., ten years after the election) since the Census updates the establishment level employment and survival records at least every five years.

Sample Selection and Matching Elections to Census Establishments

Before matching the election data to the Census data, we impose several sample restrictions to focus on certification elections likely to shift an establishment’s union status. Appendix Table A1 shows how these and subsequent sample restrictions affect the number of elections and eligible voters in our final sample. First, we restrict the sample to elections held between 1981-2005. Since the LBD starts in 1976 and ends in 2015, this gives us a five-year pre-period and ten-year post-period for all elections. Second, we drop non-representation election cases (e.g., decertification elections). Third, we drop contested elections, which are elections with multiple unions on the ballot. These elections often involve incumbent unions (e.g., “union raids”) and consequently may not be associated with changes in union representation (Sandver and Ready, 1998). Fourth, we drop elections with fewer than six workers in the bargaining unit to ensure that the election could lead to a sizable increase in union representation.

18Since we define the treatment date as the day the election petition was filed, we would expect larger effects on “year 0” payroll than employment due to these differences in the measures’ timing.
After these sample restrictions, we implement a name and address-based procedure to link each establishment-level NLRB election to a unique establishment in the LBD (our matching strategy is similar to Kline et al. (2019)). We match each election to the universe of Census establishments by calculating a weighted average of the Soft TF-IDF distance between establishment names and the geographic distance between their addresses. For each election, we select the Census match with the highest match score that surpasses a minimum threshold. This procedure yields a match for 70 % of the union elections that meet the previously described sample criteria. Additionally, we apply the same procedure to link each FMCS notice to an LBD establishment. See Appendix Section C for more details on our matching algorithm.

We further restrict the election sample to fit the requirements for our empirical strategy. For each establishment, we only keep the first election matched to that establishment. As discussed in Section 4, this means that our treatment should be interpreted as the effects of winning versus losing the first union election at an establishment. Next, we drop elections at establishments less than three years old. Since a key test of the identifying assumptions is that the outcomes for winning and losing elections evolved similarly before the election, we do not want to include observations where we cannot evaluate this for at least three time periods. Consequently, our treatment effects may not be representative of the effects of unionization at very young establishments. Finally, to keep our sample the same across model specifications, we require that each observation have non-zero payroll and employment one year before the election so that we can define outcome variables relative to then for all observations. This leaves us with an overall sample of approximately 27,000 elections (see Appendix Table A1). Finally, for much of our analysis that compares the winning versus losing elections, we restrict the sample to only include elections with vote shares between 20 and 80 % inclusive. The motivation for this inclusion is that some of the tests of our identifying assumptions discussed in Section 4 fail for the extreme vote-share elections. Appendix Table A1 shows that this further restricts our sample to 19,000 elections. We show that our main results are qualitatively robust, but less powerful, including a 30-70 % bandwidth.

Table 1 presents summary statistics for our sample of union elections. The statistics are presented separately for winning and losing elections and confirm the patterns of selection into winning elections described in Section 2. In particular, we find that winning elections are at establishments that are, on average, smaller, less likely to be part of multi-establishment firms, and more likely to have another bargaining unit at that establishment already unionized. The differences, however, are less striking for workers’ average wages or establishment age.

4 Empirical Strategy and Identifying Assumption

Our research design combines standard difference-in-differences (DiD) techniques with tests of our identifying assumption from the regression discontinuity extrapolation literature. Our identifying assumption is a conditional parallel trends assumption between elections with different vote shares.

19 Although matching introduces measurement error in our binary treatment variable, such measurement error should bias us against finding effects of unionization (see e.g., Card (1996) for measurement error in individual-level union status).
Since we observe vote shares that determine treatment assignment, we can assess several testable implications of this assumption that are not possible in a standard DiD setting.

4.1 DiD Specification and Identifying Assumptions

Potential Outcomes To fix ideas, consider just establishments, indexed by $i$, that held a union election in one year, $E_i$ (e.g., all elections in 1995). We refer to these elections as cohort $E_i$. Treatment at time $t$, $D_{it}$, is defined as both holding an election and the union receiving a vote share, $V_i$, of more than 50%\(^{20}\)

$$D_{it} = 1[V_i > .5 & t \geq E_i]. \tag{1}$$

An establishment’s non-unionized potential outcome is $Y_{it}^0$. Its unionized potential outcome is $Y_{it}^E(V)$ which depends on its cohort $E$ and election vote share $V$. This allows for dynamic treatment effects and heterogeneous treatment effects by vote share respectively. We assume no anticipation before the year of the election (i.e., $Y_{it}^E(V) = Y_{it}^0$ for all $t < E_i$). Observed outcomes are thus\(^{21}\)

$$Y_{it} = Y_{it}^0 + D_{it} \left(Y_{it}^E(V_i) - Y_{it}^0\right). \tag{2}$$

Our estimand of interest is the treatment effect $n$ years after a successful union election with vote share $V$, which is:

$$\delta_n(V) = E \left[Y_{it}^E(V_i) - Y_{it}^0 | V_i = V \& t - E_i = n \right]. \tag{3}$$

DiD Specifications For a single cohort, we can estimate the following specification

$$Y_{it} = \gamma_i + \alpha_t + \sum_n \delta_n \cdot 1[t - E_i = n] \times 1[V_i > .5] + X_i’\beta_n + \varepsilon_{it} \tag{4}$$

where $\gamma_i$ are establishment fixed effects (FEs) and $\alpha_t$ are year FEs.\(^{22}\) The coefficients of interest, $\delta_n$, represent the average, dynamic treatment effect of a successful union election. $X_i$ are baseline, one year before the election, establishment characteristics whose coefficients vary with event time $n$ (i.e., flexible trends by baseline establishment size).

Identifying Assumption Our identifying assumption is a conditional parallel trends by vote share assumption. Specifically, we assume that outcomes at establishments with different election

---

\(^{20}\)This definition assumes that treatment is absorbing (i.e., $D_{it} = 1 \Rightarrow D_{i’t} = 1 \forall t’ > t$). This assumption ignores that workers may lose union representation through a decertification election. Additionally, after losing an election, unions may hold another election. Since we only include the first election at each establishment, we interpret treatment as the dynamic effects of winning a first union election which does not correspond one-to-one with future representation.

\(^{21}\)Here, we assume that losing elections have no causal effect. This assumption is stronger than what we make in our empirical approach since we cannot disentangle the effect of losing an election from the selection into holding an election. We make this assumption for simplicity, but we could also index losing election potential outcomes by cohort to relax the assumption.

\(^{22}\)We exclude establishment FEs for outcomes that are identical for all establishments in the baseline year, $t - E_i = -1$ (e.g., establishment survival and DHS growth rates). We include them for log outcomes. See the outcome discussion for more details.
vote shares but the same baseline characteristics would have followed parallel trends had no election occurred

\[
E \left[ Y^0_{it} - Y^0_{it-1} | X_i, V_i \right] = E \left[ Y^0_d - Y^0_{d-1} | X_i \right]. \tag{5}
\]

There are multiple things to note about this assumption. First, it does not restrict selection into union elections (e.g., organizers targeting more productive establishments) or selection on gains based on the effects of unionization (e.g., workers only voting for effective unions). Second, the assumption is stronger than the standard DiD assumption because it requires parallel trends by vote share instead of only, on average, between the treated and control observations. This yields a richer set of testable implications discussed below. Third, the assumption imposes a functional form restriction on potential outcomes (Kahn-Lang and Lang, 2020; Roth and Sant’Anna, 2021) which we discuss for each specific outcomes below. Finally, as discussed in Section 2, vote shares are influenced by workers, employers, and other factors that could violate this assumption. This possibility motivates our conditioning on particular baseline \(X_i\)'s and assessing multiple testable implications of this assumption to provide reassurance that such selection is not biasing our results.

Our empirical strategy also addresses the concern that vote-share manipulation around the 50 % threshold could violate assumption 5 because elections just around the threshold are only a small share of our overall sample. For example, our vote-share heterogeneity estimates show that excluding elections right around the 50 % threshold would not qualitatively change our results.

**Testable Implications of the Identifying Assumption**

Our identifying assumption yields several testable implications. The intuition for these tests is that we observe \(Y^0_d\) for many observations and can test whether equation 5 holds for different subsets of these observations.

The first testable implication is that we should see conditional parallel trends in *pre-election outcomes across all vote shares*

\[
E \left[ Y_{it} - Y_{it-1} | X_i, V_i \right] = E \left[ Y_{it} - Y_{it-1} | X_i \right] \text{ for all } t < E_i. \tag{6}
\]

This test nests the standard DiD pre-trends test between all winning versus losing elections. Additionally, we can test for similar pre-trends between finer vote share groups. For example, we can estimate whether establishments where the union won by different margins of victory grew at different rates before the election by comparing pre-trend estimates for 50-60 % versus 60-70 % elections. This test mirrors the tests proposed by Angrist and Rokkanen (2015) and Bennett (2020) for regression discontinuity identification away from the threshold. They argue that conditionally mean independence of potential outcomes and the running variable for a given bandwidth around the RD threshold is support for being able to estimate treatment effects within that bandwidth.

One reason that we only include 20-80 % vote-share elections in our preferred specification is that we sometimes find violations of Equation 6 for extreme parts of the vote-share distribution.

The second testable implication is that we should see conditional parallel trends in *post-election*
outcomes between losing elections with different vote shares

\[ E \left[ Y_{it} - Y_{it-1} \mid X_i, V_i \right] = E \left[ Y_{it} - Y_{it-1} \mid X_i \right] \text{ for all } t \geq E_i \text{ & } V_i \leq .5. \] (7)

To implement this test, we can estimate whether the post-election outcomes for elections that lost by different margins of victory are different (e.g., compare the conditional post-election survival rates for 30-40 % versus 40-50 % elections). This test gives us one way to address the concern that election vote shares are correlated with future productivity shocks. If this were the case, we would also expect these shocks to cause differences between the outcomes at losing elections with different vote shares.\(^{23}\)

Figure 1 illustrates our identifying assumption and these testable implications. It plots average outcomes two years before the election, \(Y_{i-2}\) and \(Y_{i-1}\), and one year afterward, \(Y_{i,1}\), by vote share. Testing parallel pre-trends by vote share corresponds to comparing the distance between \(Y_{i-2}\) and \(Y_{i-1}\). Testing parallel post-trends for losing elections corresponds to comparing the distance between \(Y_{i-2}\) and \(Y_{i,1}\) for losing elections.

**Estimating Effects for Multiple Cohorts** Our sample includes all election cohorts from 1981-2005. To estimate the effect across all cohorts, we pool these elections and estimate

\[ Y_{it} = \gamma_i + \alpha_{t,E_i} + \sum_n \delta_n \cdot 1[t - E_i = n] \times 1[V_i > .5] + X'_i \beta_{n,E_i} + \varepsilon_{it}. \] (8)

This specification is the same as equation 4, except that the year FEs and baseline controls can now vary by cohort (i.e., \(\alpha_{t,E_i}\) and \(\beta_{n,E_i}\) have \(E_i\) subscripts). The motivation for this flexibility is that with cohort-specific controls, our estimates are the same as estimating \(\delta_n\) cohort-by-cohort except we use regression weights to aggregate the estimates. Consequently, there are two differences between our setting and the standard “staggered adoption” setting. First, we avoid the potential negative weight issues that arise from heterogeneous cohort-specific treatment effects (Sun and Abraham, 2020; Goodman-Bacon, 2021; de Chaisemartin and D’Haultfoeuille, 2020).\(^{24}\) Second, we only need to assume that our identifying assumption in equation 5 holds within each cohort.\(^{25}\) Both differences are because our estimates come from comparing winning and losing elections within the same cohort rather than across cohorts which might lead to negative weights or alternative parallel trend assumptions. Finally, we cluster standard errors at the firm level.\(^{26}\)

\(^{23}\)Additionally, this test allows us to evaluate one version of the “union threat” hypothesis. In particular, it allows us to test whether losing a union election by a small margin of victory affects an establishment differently than losing by a larger margin. This test, however, would not capture across-the-board union threat effects that don’t vary by vote shares.

\(^{24}\)We test for negative weights on each cohort treatment effects using Sun and Abraham (2020)’s `eventstudyweights` package.

\(^{25}\)With multiple cohorts, our identifying assumption is \(E \left[ Y_{it} - Y_{it-1} \mid X_i, E_i, V_i \right] = E \left[ Y_{it} - Y_{it-1} \mid X_i, E_i \right]\). Thus, we do not require that selection into elections in the 1980s is the same as selection into elections in the 2000s.

\(^{26}\)This accounts for serially correlated establishment-level outcomes across time and across elections at different establishments within the same firm. Our regression weighting to aggregate the \(\delta_n\) estimates easily accommodates this level of clustering.
Establishment-Level Controls  To account for observable determinants of election outcomes, we control for progressively richer establishment-level characteristics. All controls are from one year before the election and interacted with event time. The event-time interaction allows for flexible pre- and post-election trends by baseline characteristics (e.g., differential employment growth rates for large versus small establishments). Our first industry and employment controls specification, includes baseline employment and three-digit NAICS industry-by-year controls. The motivation for starting with these covariates is that they are among the strongest predictors of union election victory (see Section 2), and they are key determinants of establishment growth and survival dynamics (Dunne et al., 1989; Haltiwanger et al., 2013). Next, we add other characteristics in the LBD (baseline payroll, establishment age, and single/multi-establishment status) and an indicator for whether we observe a previous FMCS union contract at the establishment (i.e., another bargaining unit already unionized at the establishment). We refer to this specification as the pooled controls specification. Finally, we interact all controls from the previous specification with cohort (i.e., year of election). This is our preferred flexible controls specification. The cohort interactions result in the appealing within-cohort identification assumption discussed above. We show, however, that our main results are robust to pooling controls across cohorts or only including the employment and industry controls.

Establishment-Level Outcomes  The first outcome we consider is the Davis, Haltwanger and Schuh (1996) (DHS) symmetric growth rate for employment and payroll (i.e., total wage bill)

\[ G_{it} = 2 \times \frac{Y_{i,t} - Y_{i,t=E_i-1}}{Y_{i,t} + Y_{i,t=E_i-1}}. \]  

(9)

This growth rate is a second-order approximation of the log difference from time \( t \) to one year before the union election, \( E_i - 1 \). Yet, it accommodates establishment exit as \( G_{it} \) equals \(-2\) for establishments that do not exist (i.e., have zero employment). Consequently, a \(-0.2\) value of \( G_{it} \) could represent either an approximately 20% decline in intensive margin employment with no survival effects or a 10 percentage point decrease in the likelihood of survival. Since the growth rate accommodates exit, we can simultaneously evaluate pre-trends and interpret treatment effects even with treatment effects on establishment exit that could lead to a selected group of survivors. For this reason, the DHS growth rate is commonly used to analyze firm growth dynamics.

---

27Our baseline specification interacts industry by year and event time because some of our outcomes are cumulative measures (e.g., the DHS growth rates and survival). Thus, industry-by-year FEs would capture industry growth rates over different time horizons. For all continuous variables, we flexibly parameterize their functional form with decile fixed effects.

28The motivation for including the previous contract control is that union elections are more successful when other workers at the same establishment are already unionized. The selection into such elections may also differ from the selection into elections for an establishment’s first bargaining unit. When we pool all industries together, we interact controls in this specification with three coarse industry groups (e.g., manufacturing, services, and “other”). This keeps them at the same level of granularity for our overall and manufacturing estimates.

29Conventionally, the growth rate is defined annually (e.g., from \( t - 1 \) to \( t \)) but we define it over longer time-horizons to measure cumulative changes. Additionally, since our sample restrictions impose non-zero employment at \( t = E_i - 1 \), \( G_{it} \) is never equal to 2 which it usually equals for entrants. Establishments that do not exist at time \( t \) before the election have \( G_{it} = -2 \).

30See Haltiwanger et al. (2013); Chodorow-Reich (2014) for general use and Arnold (2019); Davis et al. (2014) for DiD contexts.
To estimate the effect of unionization on extensive margin employment growth, we include establishment survival as an outcome (an indicator for whether the establishment exists at time \( t \)).

We can compare the survival effect to the DHS growth rate effect to answer how much of the DHS growth rate effect is *mechanically* due to exit (e.g., \( G_{it} = -0.2 \) could be completely explained by a 10 pct. pt. decrease in survival). However, the residual, the part of \( G_{it} \) unexplained by exit, could be either intensive-margin employment changes or selective exit based on employment growth rates.\(^{31}\)

Finally, we define the outcome as log employment or log payroll. A challenge with logged outcomes is that treatment effects on establishment survival can bias comparisons of potentially selected *survivors*. The pre-trends for these log outcomes, however, are a useful complement to the DHS growth rate pre-trends.\(^{32}\) For interpreting the treatment effects on these log outcomes, we provide two ways of partially alleviating the selective exit concern. First, all specifications with log outcomes include establishment FEs that account for *level* differences between the surviving and exiting establishments.\(^{33}\) Second, for some results, we argue that the timing of the effects on logged variables versus the effects on survival suggests intensive margin effects. Yet, we still recommend interpreting the treatment effects for these log outcomes with caution since we cannot completely eliminate potential bias from selective exit.

For each outcome, we make a different parallel trends functional form assumption. For log outcomes, we view the restriction that log employment and payroll would have (conditionally) evolved in parallel as a reasonable restriction in this setting.\(^{34}\) Additionally, we can test whether the restriction holds in the pre-period (i.e., equation 6). For establishment survival, we assume that the survival probabilities between elections with different vote shares would have (conditionally) been equal had no election occurred at the establishments. For this assumption, we cannot test whether the parallel trends assumption holds in the pre-period since all establishments exist at event-time zero. However, we can test whether this functional form assumption holds between the losing elections with different vote shares. For the DHS growth rate, the outcome is approximately a linear combination of the log employment and survival probabilities so parallel trends in both other outcomes imply parallel trends in the DHS growth rate.\(^{35}\)

### 5 Empirical Results: Overall Employment and Survival Effects

In this section, we examine the relative effects of successful unionization on establishment employment and survival. We first analyze the overall differences in employment growth rates between

---

\(^{31}\)For example, larger exit effects on faster-growing establishments would create a residual unexplained by exit.

\(^{32}\)In theory, the DHS pre-trends combine intensive and extensive margin employment changes. However, in specifications where we control for baseline establishment age, the DHS pre-trends will closely approximate pre-trends for log outcomes.

\(^{33}\)For DHS growth rates and establishment survival we do not include establishment fixed effects. For DHS growth rates we already capture the time-invariant establishment component by differencing relative to \( t = E_i - 1 \). For survival, it is unclear what time-invariant characteristic the fixed effect would capture.

\(^{34}\)For example, consider two firms with the same Cobb Douglas production function parameters but different baseline TFP and/or input and output prices. In response to the same demand shock (e.g., the same proportional change in the price of output), their log output and log employment would both evolve in parallel while their levels would diverge.

\(^{35}\)Specifically, we assume \( E[\Delta \ln Y^0_{it} | X_i, V_i] = E[\Delta \ln Y^0_{it} | X_i] \) and \( E[1\{Y^0_{i,t} = 0\}|X_i, V_i] = E[1\{Y^0_{i,t} = 0\}|X_i] \) which imply \( E[G^0_{it} | X_i, V_i] \approx E[G^0_{it} | X_i] \).
establishments with winning and losing elections. Next, we implement the multiple tests of our parallel trends identifying assumption described in Section 4. Finally, since we later focus on manufacturing, we present our estimates and falsification checks separately for all industries pooled together and just for elections in manufacturing.

5.1 Overall Employment and Survival Estimates

We start by estimating establishment employment growth around successful versus unsuccessful elections. Figure 3 plots the $\delta_n$ coefficients from estimating the “pooled cohort” specification in equation 8 for union elections with 20-80 % vote shares. Panel A. plots the estimates for DHS employment growth relative to one year before the election (time $E_i - 1$). Panel B. instead includes log employment as the outcome. Both panels include estimates with no controls (e.g., only year by cohort fixed effects), the industry and employment control, and the flexible control specification described in Section 4.

The results in Figure 3 show that, conditional on baseline industry and employment, establishments with successful elections had similar pre-election growth rates to establishments with unsuccessful elections but experienced large relative employment decreases following the election. The first, “no control” estimates show that establishments where the union won had relatively slower pre-election employment growth rates compared to establishments where the union lost. However, the following specification that conditions on just baseline employment and industry yields similar pre-election growth rates for DHS and log employment. Starting one year after the election, this specification shows decreased employment for establishments with successful union elections. The effects stabilize approximately three years after the election. Finally, the results from the third flexible control specification are very similar to just including industry and employment controls.

Panel C. in Figure 3 adds estimates for DHS payroll growth and establishment survival from the flexible control specification. The payroll estimates show a faster initial decrease compared to employment. This difference could be due to either compositional shifts to low-wage workers or differences in the timing of the payroll versus employment measures described in Section 3. Five years after a successful union election, the cumulative DHS employment and payroll growth rates are -0.13 and -0.14 lower than establishments with unsuccessful elections (consistent with a 14 % decrease in payroll or a 7 pct. pt. decrease in survival likelihood). Appendix Figure A3 presents estimates from the same specification for log employment and payroll. These estimates allow us to reject five-year, pre-election growth rate differences of more than 3.5 % for employment and 1.8 % for payroll. Unlike the DHS measures, we find larger five-year log payroll than employment declines. Although this evidence would be consistent with long-run compositional changes, we recommend interpreting it cautiously given potential biases from selective exit.

The survival estimates in Panel C. indicate that most of the decrease in DHS employment

---

36 Without any controls, the DHS growth rates and log employment pre-trends measure somewhat different growth rates. The DHS employment growth rates combine intensive and extensive margin changes, while log employment only captures intensive margin changes. In the control specifications that include establishment age, the measures should approximate the same concept.
and payroll growth rates is from a lower likelihood of establishment survival. To decompose what share of the DHS effects is from survival effects, we plot the survival estimates on a separate y-axis scaled to be one-half the DHS growth rate axis. Comparing the exit and DHS coefficients illustrates how much of the DHS effect can be mechanically explained by the survival effect (see Section 4). The estimates show that five years after an election, establishments with successful elections are four pct. pts. less likely to survive, and this effect increases slightly to five pct. pts. after ten years. Consequently, about two-thirds of the -.13 five-year DHS employment growth rate estimate is mechanically due to decreased establishment survival. Finally, the relatively slower timing of the survival versus employment effects is consistent with legal considerations about closing establishment following union elections. For example, Munger et al. (1988) describe how a short time between an election and establishment closure could be evidence that the closure was an unfair labor practice with the “intent to chill unionism” across the entire firm.

Given our later focus on manufacturing, Figure 4 presents analogous results including only manufacturing elections. Overall, the results mirror the results for all industries pooled together but there are two exceptions. First, we find similar pre-election employment growth rates even with no baseline controls. Second, the magnitude of the effects is larger than the effects for all industries (e.g., the five-year DHS employment estimates are -0.17 versus -0.13, and the five-year survival effects are -0.05 versus -0.04, respectively). We show later that this difference is because the effects of unionization in the service industries are generally small.

5.2 Vote-Share Heterogeneity Tests of Identifying Assumption

We next provide further evidence that our results are driven by unionization by assessing several testable implications of our identifying assumption. Additionally, we estimate treatment effect heterogeneity by the unions’ margin of support. To implement these tests, we first present visual evidence of how treatment effects and pre-election trends vary across the vote-share distribution and then formally test for linear trends by vote share.

Nonparametric Vote-Share Heterogeneity  To estimate pre-trends and treatment effects for different parts of the vote-share distribution, we estimate the following modified version of our main DiD specification

\[
Y_{it} = \alpha_{t,E_i} + \sum_{g} \sum_{n} \delta_{g,n} \cdot 1[t - E_i = n] \times 1[V_i \in V^g] + X_i'\beta_{n,E_i} + \epsilon_{it} \tag{10}
\]

where \(V^g\) are exhaustive subsets of the vote-share distribution. We partition the vote-share distribution into eight groups (0 – 20 %, 20 – 30 %, 30 – 40 %, 40 – 50 %, 50 – 60 %, 60 – 70 %, 70 – 80 %, and 80 – 100 %). We omit the 20 – 30 % group so the estimates for each group are relative to outcomes for 20 – 30 % vote share elections. This specification allows us to assess the two testable implications of our identifying assumption described in Section 4. First, we test whether establishments’ pre-election outcomes are similar across the vote share distribution by comparing...
\(\delta_{g,n}\) estimates for \(n < 0\) (i.e., testing equation 6). Second, we test whether post-election outcomes differ between losing elections with different vote shares by comparing \(\delta_{g,n}\) estimates for \(n > 0\) and \(\nu^g \in [0, 50]\) (i.e., testing equation 7). For presenting these estimates, we start with manufacturing where the results closely support our identifying assumption, making the setup easier to explain. We then turn to estimates for all industries where there is evidence of rejections of these tests for some outcomes. We find, however, that the violations are driven by elections that lost by exactly 50 % which we would expect such differences.

Figure 5 presents results from estimating equation 10 for all manufacturing elections. The estimates include our flexible controls specification (see the following vote-share heterogeneity analysis for robustness to alternative controls). Panel A. includes pre-period and treatment-effect estimates for each vote share group with DHS employment growth as the outcome. First, the five-, three-, and two-year pre-trend estimates are similar across almost the entire vote-share distribution relative to 20-30 % elections (the one exception is 0-20 % elections which are excluded from our main sample). These results provide support for our identifying assumption by showing that untreated potential outcomes (here pre-election outcomes) are mean independent of \(V_i\) conditional on \(X_i\). Second, none of the five- and ten-year treatment effect estimates for losing elections are significantly different than the estimates for 20-30 % elections. These results similarly test equation 5 but instead use post-election untreated potential outcomes. They also provide one reassurance against the concern that our results are driven by future productivity shocks correlated with vote shares. In that case, we would also expect such shocks to cause different outcomes for losing elections with different vote shares. Finally, the five- and ten-year treatment effect estimates increase in the union vote share but are not statistically different (e.g., -0.18 versus -0.28 ten-year estimates for 50-60 % and 70-80 % elections, respectively).

Figure 5, Panel B. plots the same estimates with establishment survival as the outcome variable. Although we cannot test for pre-trends in establishment exit rates, we can test our parallel trends assumption for establishment survival by testing whether losing elections’ with different vote shares had different post-election survival rates. Reassuringly, the survival rates for all losing election vote-share groups are not statistically different than the survival rate for 20-30 % elections. For winning elections, however, the figure shows that the long-run effects on business survival increase in the union vote-share although the differences are not statistically different across groups.

Figure 6 presents the vote-share heterogeneity estimates for elections in all industries. Panel A. shows that for our main sample of 20-80 % vote-share elections, we find very similar pre-election employment growth rates across the vote-share distribution. For 0-20 and 80-100 % vote share elections, however, we find evidence of different pre- and post-election growth rates which is one motivation for excluding these elections from our main analysis. For post-election outcomes, we find similar DHS employment growth rates between 20-30 % and 30-40 % elections but find somewhat negative estimates for 40-50 % vote share elections. The ten-year estimate for 40-50 % elections is also significantly different from zero at the 10 % level. However, these negative estimates are driven by elections where the union received exactly 50 % of votes, and there are multiple reasons to expect
differences between 50% vote share elections and elections where the union lost by slightly larger margins.\footnote{There are multiple potential reasons for outcome differences at establishments with 50% vote-share elections. First, due to the discreetness of total votes, elections with exactly 50% vote shares have a small number of total votes cast (see the “integer problem” in DiNardo and Lee (2004)). For example, based on the NLRB data, the median (mean) number of voters in 50% vote-share elections is 12 (22) compared to 50 (96) voters in elections with vote shares in the [45, 50) range. Although our employment controls capture corresponding establishment size differences, they do not capture potential differences in the bargaining unit size to employment shares. Second, the manipulation around the 50% threshold is largely due to challenges to single votes which disproportionately affects elections with 50% vote shares Frandsen (2017). Finally, experiencing an election where the union loses by just one vote may have a different effect on employers than losing by larger margins of support.} To see this, when we estimate the 40-50% effects excluding elections 50% vote shares, the five-year estimate is -0.015 (std. err. 0.025) and the ten-year estimate is -0.032 (std. err. 0.028). Both estimates are much smaller than the treatment-effect estimates for the neighboring group of winning elections with 50-60% of votes (-0.11 and -0.16 at the two time horizons). Additionally, Panel B. of Figure 6 shows that for exit rates, there is no evidence of differential survival rates between 20 − 30, 30 − 40, and 40 − 50% losing elections in the sample.

**Parametric Vote-Share Heterogeneity** To complement the previous non-parametric analysis, we estimate a series of parametric vote-share heterogeneity tests. Specifically, we test for linear trends in pre- or post-election outcomes by vote share. There are two motivations for this extension. First, these tests may have more power. Second, they provide a parsimonious way to assess robustness to different sets of controls. We show that our estimates are qualitatively the same only including the more limited employment and industry controls.

We first test for a linear trend in pre-election employment growth rates across the entire votes-share distribution. Second, we test for linear trends in post-election outcomes separately for winning and losing elections. To implement these tests, we estimate a modified version of the specification in Equation 8. Specifically, rather than only interact event-time with the winning indicator (e.g., \(1[V_i > .5]\)), we instead include the following interactions with event-time\footnote{\(V_i - .5\) is equal to \((V_i - .5) \times 1[V_i > .5]\). Since we only estimate this specification for elections with 20-80% vote shares, we actually shift the vote-share variables to all start at zero (e.g., subtracting .2 from the \(V_i\) variables and .3 from the winning vote share variable). This ensures that the vote-share coefficients only capture slope and not level differences.}

\[
\begin{align*}
1[t - E_i = n] \times \begin{cases} 
\rho \cdot V_i & \text{if } n < 0 \\
\eta \cdot 1[V_i > .5] + \theta \cdot V_i + \tau \cdot [V_i - .5]^+ & \text{if } n \geq 0 
\end{cases}
\end{align*}
\] (11)

For the pre-trend vote-share heterogeneity test (i.e., \(n < 0\)), we do not include an interaction with treatment so pre-period “treatment effects” could be captured in the \(\rho\) estimates. For the post-election outcome tests, we include an interaction with treatment so \(\eta\) that captures the treatment effect for 51% elections (e.g., a linear RD estimate). Consequently, the \(\theta\) and \(\tau\) coefficients estimate slope differences that do not include the effect around the 50% threshold.

Table 2 includes estimates of pre-election growth rate trends by vote share, \(\rho\), for one to five years before each election. The estimates are for the main 20-80% vote share sample and are presented separately for all industry and manufacturing elections. We present estimates with the employment and industry and flexible control specification. Across all estimates, we never find
significant pre-election growth rate trends.\textsuperscript{39} We view this as strong evidence that establishments’ conditional pre-election employment growth rates are similar across the vote share distribution.

Table 3 presents the estimates testing for post-election outcome vote-share trends. We present estimates of separate slopes for losing elections (i.e., $\theta$) and winning elections (i.e., $\theta + \tau$). These tables include our preferred flexible control specification but Appendix Table A2 shows qualitatively similar results with the employment and industry controls. Motivated by the potential issues with 50 \% vote share elections (see above discussion), we also present the estimates separately with and without excluding 50 \% elections from the estimates.

The results for all industries in Table 3, Panel A. indicate significant negative trends in DHS employment growth rates by vote share for both losing and winning elections. However, mirroring the non-parametric analysis, when we exclude the 50 \% elections, we do not detect significant trends for losing elections. We do, however, find negative significant trends by vote share for winning elections, consistent with increasing treatment effects for larger margin of support elections. For example, for five-year DHS employment, we estimate a vote-share trend of -0.066 (std. err. 0.122) for losing elections and -0.389 (std. err. 0.149) for winning elections. For establishment survival, we never find significant trends for winning or losing elections. For some specifications, the losing election trends are actually positive providing support that our overall survival estimates are not driven by a negative correlation between election vote shares and establishment survival.

The manufacturing estimates presented in Table 3, Panel B. are similar to the estimates for all industries pooled together. Without excluding the 50 \% elections, we find negative although insignificant DHS employment trends for losing elections. However, dropping the 50 \% elections results in smaller trends for losing elections and large although insignificant vote-share heterogeneity estimates for winning elections (e.g., five-year DHS trend estimates of -0.072 (std. err. 0.199) for losing elections and -0.406 (std. err. 0.299) for winning elections). Using the benchmark described in footnote \textsuperscript{39}, the five-year estimate in Column (2) could explain -0.036 of the -0.195 DHS employment growth rate treatment effect estimate for the 70-80 \% elections and the ten-year estimate could explain -0.105 of the -0.277 total effect. The manufacturing establishment survival estimates are also never significant for winning or losing elections and, at times, positive for losing elections.

**Employment and Survival Effect Robustness** We next present two additional checks further validating our overall estimates of the negative impacts of unionization on establishment survival and employment. First, we assess whether our estimated effects increase in the size of the bargaining unit (Lee and Mas (2012) conduct a similar test). The motivation is that the relative share of unionized workers should mediate many direct effects of unionization but not some potential violations of our identifying assumption (e.g., workers voting based on their expectations of future company performance or managers’ competence). Appendix Table A3 presents the coefficient estimates from

\textsuperscript{39} To assess the magnitude of the estimates, the largest positive point estimate is .05. A reasonable benchmark is what the estimates imply for the differences between 20-30 and 70 -80 \% elections presented in Figure 6. Since the midpoints between those bins are .5 apart, the .05 coefficient implies a small difference in pre-election employment growth rates of around 2.5 \% between 20-30 and 70 -80 \% elections. We note, however, that the 95 \% confidence intervals on some of the estimates would include relatively large pre-election growth rate differences.
interacting the three-, five-, and ten-year treatment indicators with the share of the establishments’ total employment covered in the bargaining unit (see Appendix Section C for details). It shows that the three and five-year treatment effects are significantly increasing in the bargaining unit share for both outcomes confirming that the effect seems mediated by the share of workers gaining representation. The interactions, however, are no longer significantly different than zero at the ten-year horizon. One explanation for the lack of persistence is that the relative size of the bargaining unit versus establishment employment could change substantially over time.

Second, Appendix Figure A4 plots DHS employment growth rate estimates with ten-year pre- and post-periods.\textsuperscript{40} First, it shows no evidence of large pre-trends in employment growth rates up to ten years before elections both in manufacturing and for all industries pooled together. Although for all industries, we find significant pre-period estimates 6, 7, and 8 years before the election, they are economically small (e.g., approximately 1.7 to 2.0 percent differences). Moreover, the ten-year pre-period estimate is insignificant and it’s confidence interval allows us to rule out employment growth differences of more than approximately 3.2 percent. Second, the figure shows that the post-election effects are relatively stable starting three years after the election. For manufacturing, however, there is a slight increase in the effect from years 5 to 10.

5.3 Industry-Specific Employment and Survival Estimates

We next separately estimate the effects for different industries and show that the overall effects are driven by non-service-sector union elections. There are multiple reasons to expect substantial heterogeneity across industries. First, the quality of post-election labor relations may differ across sectors (e.g., the higher post-election strike propensity for manufacturing elections in Figure 2 Panel D. suggests more adversarial relations). Second, firms in different industries may differ in how easily they can “avoid unionization”. For example, mobile, multi-establishment manufacturing firms may avoid working with new unions by shifting production to non-unionized plants. However, this tactic may be difficult in non-tradable industries (e.g., hospitals) or in tradable industries with ties to their local area (e.g., hotels).

To estimate this heterogeneity, we classify our elections into three broad industry groups – manufacturing, services, and a residual “other” group.\textsuperscript{41} Weighted by the number of eligible voters, 70 % of our service-sector elections are for healthcare (e.g., hospitals and nursing homes), security,

\textsuperscript{40}Note, since our data start in 1976 and our elections start in 1981, the -10 to -6 estimates are from an unbalanced panel (e.g., the -6 point estimate is based on a different number of election years than the -5 estimate, etc.) This is one motivation for why we focus on the -5 to 10-year estimates with the balanced panel for the main analysis.

\textsuperscript{41}These classifications are based on the Fort and Klimek (2016) 2012 NAICS sectors of the Census establishment we match to each election. We define manufacturing as NAICS sectors 31-33 and services as NAICS 51-81 and NAICS 44-45. Our classification of services differs slightly from other measures. For example, compared to the sampling frame for the Census’s Service Annual Survey, we include retail trade in the services group and exclude utilities and transportation and warehousing. The motivation for these changes is that we want to capture a notion of “service-sector unionization.” Retail workers (e.g., grocery store workers) are commonly referred to as part of service sector unionization. As evidence of this, the “OUR Walmart” campaign was frequently described as attempting to unionize service workers (Brown, 2011). Alternatively, most elections in utilities involved electrical workers (more similar to crafts unions in the building trades) and many elections in transportation and warehousing involved drivers.
restaurants, grocery stores, universities, and print media establishments. The other category includes agriculture, construction, mining, transportation and warehousing, utilities, and wholesale trade.

To estimate the industry-specific heterogeneity, we use the following specification for a discrete heterogeneity variable $H_i$ (e.g., the three industry groups):\footnote{This specification has two advantages relative to restricting the sample for each value of $H_i$. First, we can pool the control coefficients across heterogeneity groups and use all the data to estimate their coefficients. For all heterogeneity estimates, we also add the specific heterogeneity group as an additional control in $X_i$ so that we account for any differential trends by the specific heterogeneity groups. Second, it allows for Wald tests of equality across the different heterogeneity groups.}

$$Y_{it} = \alpha_{t,E_i} + \sum_h \sum_n \delta_{h,n} \cdot [t - E_i = n] \times [V_i > .5] \times [H_i = h] + X_i' \beta_{n,E_i} + \epsilon_{it}. \quad (12)$$

The $\delta_{h,n}$ coefficients now estimate the dynamic effects of successful union elections for elections with $H_i = h$. We also estimate all subsequent heterogeneity in Section 6 using equation 12.

Table 4 presents the DHS employment growth and survival effects estimated separately for each industry group. First, there is limited evidence of employment growth pre-trends for any of the groups. The only marginally significant pre-period estimate is for the service sector where we find the smallest main effects. Second, the overall employment and survival decreases are driven by similarly sized effects for elections in the manufacturing and the other sector. For elections in the service sector, the effects of unionization are substantially smaller. For example, the five-year DHS employment growth estimates for manufacturing and services are -0.174 (std. err. 0.029) and -0.057 (std. err. 0.024), respectively. Additionally, the ten-year survival estimate for the service sector is not significantly different than zero and the confidence interval allows us to reject effects of more than four pct. pts.

Appendix Table A4 shows that the smaller estimates in the service sector hold under several alternative specifications. Specifically, it presents the point estimate and standard error of the difference between the manufacturing and service-sector coefficients over each time horizon. The effects in manufacturing remain significantly larger when we (1) pool controls across cohorts, (2) restrict the sample to 30-70% elections, and (3) restrict the sample to elections where the bargaining unit was at least 25% of establishment employment. The results with the last restriction show that the effects of successful union elections in the service sector are not smaller because these elections are more likely to include relatively smaller bargaining unit shares.

6 Testing for Manager Opposition and Union Avoidance

After documenting the large overall impacts of successful union elections on establishment employment and survival, we test whether some of this effect is due to managers’ dislike of working with unions or their ability to easily avoid working with unions. For this analysis, a sector-specific analysis is the most appropriate because the tactics that employers can use to avoid unions may differ across sectors. In manufacturing, a common union avoidance tactic for multi-establishment firms during this time was shifting production away from unionized establishments to non-unionized
establishments (Bluestone and Harrison, 1982; Verma, 1985; Kochan et al., 1986a). However, in construction, one of the industries in the “other” industry group where we also find negative effects, most firms are single-establishment firms, so they cannot shift production across establishments (Butani et al., 2005). So, the same test might not capture union avoidance across sectors. Consequently, we focus on elections in manufacturing for three reasons. First, manufacturing is the largest sector where we find negative effects. Second, as discussed above, we can use the high-quality establishment and firm linkages in our data to test for union avoidance via production shifting. Finally, for manufacturing, we have detailed measures of plant-level productivity that we use in our last test of whether unionization leads to productivity reallocation.

Multi- versus Single-Establishment Manufacturing Firms The first part of our hypothesis is that firms avoid working with unions by shifting production away from newly unionized plants to other plants. Since this shifting is only possible for firms with multiple plants, we start by estimating whether the effects of union elections are larger at establishments that are part of multi-establishment (or multi-unit, MU) firms (e.g., an election at one of the plants for a firm with several plants) versus single-establishment (SU) firms. Specifically, we define “an election at a MU firm” based on whether the election establishment’s firm had at least one other establishment under its control one year before the election.

Figure 7 plots the coefficient estimates for the DHS employment growth and survival effects for elections at SU versus MU firms. The left panel plots the cumulative DHS employment growth rates for five years before and three, five, and ten years after the election. Below each x-axis label, we include the p-value of the difference between the SU and MU estimates. Reassuringly, there is no evidence of differential pre-election employment growth rates for either group. After the election, we see significantly larger employment growth declines for elections at MUs at the three and ten-year horizon. The estimates for SUs are, however, still negative and significant. For establishment survival, the differences are even more striking. For all post-election time horizons, the effects are significantly larger for MUs, and none of the survival estimates for SUs are significantly different than zero. For example, the ten-year survival estimates are $-0.122$ (std. err. $0.021$) versus $-0.029$ (std. err. $0.029$) for MUs and SUs, respectively.

Appendix Table A5 shows the robustness of these estimates to (1) including controls pooled across cohorts and (2) only using 30-70% vote-share elections to estimate effects. It presents the difference and standard errors between the SU and MU estimates. The estimates are very similar with only pooled controls. For the 30-70% bandwidth, we still estimate substantially larger survival effects for MU establishments (e.g., six pct. pts. at the ten-year horizon) but the larger standard

---

43 Alternatively, in construction, there is evidence of firms avoiding unions by going “dual shop” and opening a new non-unionized shop that does previous work of the unionized shop (Evans and Lewis, 1989). Although we may see these new establishment openings in our data by linking establishments across Census firmids, owners may try to disguise the common ownership of these establishments to avoid potential labor-law issues with going “dual shop” (Milne, 1985). In other sectors, there is also evidence of employers using temporary workers to replace newly unionized workers (Hatton, 2014).

44 For all sectors in our manufacturing and other industry groups where we find large negative effects, manufacturing makes up 54% of elections compared the 18% for transportation and warehousing (the next largest sector). Weighting by the number of eligible voters, manufacturing comprises 68.8% of voters due to its relatively large bargaining units.
errors only lead to a significant difference for survival at the five-year horizon.

We interpret these results as showing that the effects on establishment survival in manufacturing are driven by plant closings at MU firms. For the overall employment declines, the effects are significantly larger at the multi-establishment firms but still substantial for SUs. These results indicate that MUs and SUs respond differently to unionization with MUs reducing employment primarily through plant closing while SUs respond more through intensive-margin changes. This evidence is also consistent with MU firms responding to unionization by shifting production across plants which we investigate more rigorously next. As an alternative explanation, MU firms may have a greater incentive to react strongly to unionization due to concerns about unionization spreading to their other establishments. We also investigate this later by focusing on entirely non-unionized MUs where this incentive maybe even sharper.

**Employment Shifting after Successful Elections** Next, we directly test the hypothesis that manufacturing firms avoid new unions by shifting production to other plants. Specifically, we analyze whether a successful election at one of a firm’s manufacturing plants affects the employment and survival of the firms’ other plants. While the production-shifting hypothesis predicts positive effects on other establishments, other mechanisms like input-output linkages or firms’ financial constraints predict negative spillovers (Giroud and Mueller, 2017). One prediction of the production-shifting hypothesis, however, is that the positive effects should be the largest at plants where it is easiest to produce the same products as the election plant. Consequently, we start by only considering the effects on other manufacturing establishments and then consider establishments within the same NAICS three-digit industry as the election establishment.

To construct the sample for this analysis, we start with all manufacturing elections in a specific year at multi-establishment firms. Next, we take all of the firm’s other manufacturing plants that existed during the election year and never experienced their own union election. We then calculate these plants’ DHS employment growth rates before and after the election relative to one year before the election. Finally, we stack these observations from all cohorts together and estimate a modified version of our main DiD specification. The two differences from our main specification are that (1) relative time and vote share variables are defined from the election at the firms’ other establishment, and (2) we weight the regression by each establishment’s share of its firms’ total employment. The reason for the weighting is that the sample could include multiple plants.

---

45 As discussed in Section 3, the allocation of employment across establishments at multi-establishment firms could also lead to spurious short-run negative effects (e.g., attributing some of the decreased employment at the election establishments to the firms’ other establishments).

46 LaLonde et al. (1996) similarly analyze within-firm employment spillovers of successful union elections. They do not find any evidence of spillovers but only consider the effects on all other manufacturing plants, where we also do not find only spillovers. We only find evidence of spillovers when we focus on other plants within the same three-digit NAICS industry. Bradley et al. (2017) similarly find that firms shift RD activity away from newly unionized establishments.

47 This construction results in some establishments being in the data set multiple times if their firms experience multiple union elections. For our baseline analysis, we avoided this problem by taking the first election at each establishment. However, since the Census firmids change over time even for establishments that stay in business and since establishments can move across firmids, such conditioning is more difficult here. This also motivates our two-way clustering by firm and establishment.

48 For the employment total in the denominator, we only include employment at other establishments included in the sample
matched to each election and we want to weight each election equally (i.e., not give the most weight to elections at firms with the most other plants). For this specification, we two-way cluster the standard errors by firm and establishment.\footnote{Specifically, we use each establishments’ firmid during the election year (e.g., the clustering variable is fixed over time). Yet, since an establishment could appear multiple times in the sample at different firmids, establishment-level clustering is not nested by firmid-level clustering.}

Figure 8 Panel A. plots the event-study estimates for each firms’ other establishments’ employment growth rates before and after a successful election. It presents estimates that include all manufacturing plants and estimates that only include plants in the same three-digit NAICS industry as the election establishment. For all manufacturing plants, there is no evidence of relatively higher employment growth at the other establishments following successful elections. There are two things to note about this result. First, even if firms shifted employment away from newly unionized plants it is not surprising that we do not find spillovers when we include all other plants because these plants may have produced very different products than the election plant. Second, it is reassuring that we do not estimate lower post-election employment at the other plants of firms with successful elections. If our estimates of the direct effects of unionization were biased by unrelated productivity shocks, we might expect these shocks to be at the firm level. Yet, these spillover results allow us to rule out differences in firm-level DHS employment growth rates of more than -0.04 which is much smaller than our overall estimates.\footnote{Since there may be spillovers, however, we cannot view this exercise as a formal placebo check.}

When we restrict the sample to plants that produce similar products as the election plants, we find significant employment growth spillovers. The solid estimates in Figure 8 show that the average post-election DHS employment growth rate at the firm’s other establishments in the same three-digit NAICS industry is significantly higher for firms with successful versus unsuccessful elections. Two years after the election, we estimate growth rate differences of 0.05, which persists three and four years after the election. However, the effect becomes insignificant five years after the election and at the ten-year horizon is .02 (std. err. = .03). Additionally, as shown in Table 5, some of the short-run increase in employment growth is from effects on establishment survival. Overall, the point estimates for these spillover estimates are fairly large (e.g., we estimate that a same-NAICS 3 plant is two percentage points more likely to survive three years after a successful union election at its firm compared to such plants at firms with losing elections). However, the somewhat large standard errors and the fact that the effect does not persist in the long run suggest that the true effects could be somewhat smaller.

Figure 8 Panel B. further splits up the elections based on whether or not the election establishment made up a large share of the firm’s total employment. The motivation is that when the election establishment was only a small share of the firm’s overall employment, we would not expect to have enough power to detect spillovers. We specifically split up elections based on whether the election establishment was more than 10 % of the firm’s employment in the same three-digit NAICS industry during the election year.\footnote{This heterogeneity specification is estimated the same as other heterogeneity specifications (e.g., estimated jointly with} The estimates in Panel B. show that the overall effect...
on three-digit NAICS matches is driven by relatively large elections. The fact that the effects are
driven by same industry plants and by relatively large elections is reassuring for two reasons. First,
these are the types of plants where we would expect to detect the most production shifting. Second,
it is not clear why we would also expect potential threats to our parallel trend assumptions to be
more pronounced for these specific groups.\footnote{For this analysis, the concern that would violate the parallel trends assumption is that the other plants at winning election firms are growing faster than other plants at losing election firms. Note, this is the opposite of the potential concern for our direct survival and employment result.}

Overall, we interpret these results as evidence that successful union elections lead to faster
employment growth at the firm’s other establishments. This suggests that firms shift production
away from the newly unionized plant. Interestingly, the effects also seem partially driven by
differences in survival which is consistent with some shifting being driven by decisions over which
plants to close. However, the fact that the effects do not persist in the long run and that the
estimates are somewhat imprecise, make us unable to calculate how much of the direct effects are
offset by this shifting.

**Firms’ Unionization Status** The second part of our hypothesis is that the effects of unionization
are greater when management is more opposed to the union. To test this, we estimate treatment
effect heterogeneity based on two proxies for managers’ opposition. First, we estimate effects
separately for MU firms with and without any other unionized establishments. During the time
period we study, many firms were committed to a “non-union strategy” of remaining or becoming
non-unionized (Kochan et al., 1986a). Consequently, firms without any unionized establishments may
have responded more strongly to successful unionization out of this desire to remain non-union or
due to their anti-union philosophy.\footnote{\cite{Kochan et al., 1986b} provides support that a firm’s non-unionization status represents its desire to remain non-unionized by showing that the strongest predictor of partially unionized firms’ anti-union commitment was the share of total employment unionized. One explanation for this is that a firms’ other unionized workers could pressure the firm for union avoidance tactics while less- or non-unionized firms would face less pressure. An anecdotal example of this is the failure of GM’s “southern strategy” of opening non-unionized plants in the south due to pressure from the UAW \cite{Nelson, 1996}.} Additionally, similar to Selten (1978)’s “chain store paradox”,
on-unionized firms may have a strong incentive to aggressively deter the first unionization attempt
(even if not economically profitable when considering each establishment in isolation). To test this,
we split up our elections at MUs based on whether we observe an FMCS contract at any of the
firm’s establishments in the five years before the election.\footnote{We include previous contracts at both the election establishment (i.e., a separate bargaining unit already unionized) and at all of the firm’s other establishments. See Appendix Section C for details.} Since the FMCS contract data start
in 1984, we only classify MUs as unionized versus non-unionized for elections starting in 1985 but
show robustness to starting in 1990.\footnote{As with the previous heterogeneity estimation, these estimates still include all manufacturing elections (i.e., to estimate the pooled controls), but we do not report the effects for elections at SUs or elections before 1985 from this specification.}

Figure 9 mirrors the previous heterogeneity figure and plots the DHS employment growth
rate and survival estimates for MU firms with and without unionized establishments. For overall
employment growth rates, the non-unionized firms have larger employment decreases that are
pooled controls and adding a control for the heterogeneity group by event time.
significantly larger than the unionized firms at the five- and ten-year horizons. For establishment survival, the differences are rather small and insignificant at the three- and five-year horizon. However, at the ten-year horizon, the negative survival effect is substantially larger for elections at non-unionized MU firms – -.20 (std. err. .040) versus -.09 (std. err. .027).

Appendix Table A6 shows the robustness of these differences across three changes. First, we only include pooled controls. Second, we only classify MU firms into unionized versus non-unionized firms starting in 1990. Third, we only use elections in the 30-70 % vote share bandwidth to estimate the effects. Our estimated differences are larger when we define firms’ unionization status starting in the 1990s and relatively similar for the other two specifications.

These estimates show that the long-run negative effects of unionization are substantially larger at firms without any previously unionized establishments. This evidence is consistent with these firms strategically closing or shifting employment away from the newly unionized establishments. This mirrors the documented pattern at the time of partially unionized firms shifting investment and production to non-unionized plants (Verma, 1985). Such strategic shifting could be motivated by a dislike of working with unions (e.g., Foulkes (1980) presents evidence that many non-union companies were motivated by a philosophical opposition to unions) or a fear that the unionization would spread to other plants.

**Election Delay Time** To estimate whether the negative effects of unionization are related to managers’ opposition to the union, we estimate treatment effect heterogeneity based on the delay between the election filing date and the actual election date. The motivation is that managers frequently use tactics that delay the election date to try to win the election. First, delay itself can reduce support for the union. In “Confessions of a Union Buster,” Levitt and Conrow (1993) write that the NLRA “presents endless possibilities for delays, roadblocks, and maneuvers that can undermine a union’s efforts and frustrate would-be members” and that this delay “steals momentum from a union-organizing drive, which is greatly dependent on the emotional energy of its leaders and the sense of urgency among workers.” Additionally, other tactics managers employ to influence elections (e.g., challenging the composition of the bargaining unit or filing unfair labor practice charges) add time before the election.56 Supporting this interpretation of delay time as a proxy for managers’ opposition to the union, many analyses have found that delay time is associated with lower election win rates (Roomkin and Block, 1981; Ferguson, 2008).

We start by defining election delay time and verifying that it is related to election outcomes in our sample. We define delay time as the number of days between the date the election petition was filed to the NLRB and the date the election was held (see Appendix Section C for more details). The average election delay in our sample is 62 days, and the 10th and 90th percentiles are 31 and 80 days. Appendix Figure A5 shows that our measure of delay time is negatively associated with union election success rates and positively associated with the probability of any challenged votes in the election (another proxy for managers’ anti-union campaign intensity). These relationships also

56 Another determinant of delay time is that elections with larger bargaining units have longer delays. This motivates our robustness check that the effect heterogeneity holds conditional on election unit size.
hold conditioning on other election characteristics that may be correlated with delay time (e.g., the number of eligible voters, four-digit NAICS industry, and state).

To analyze whether the negative effects of unionization differ by election delay, we start by estimating treatment effect heterogeneity separately by terciles of the within-year delay time distribution. Figure 10 plots the estimated effects for the first and third terciles for DHS employment growth (left panel) and establishment survival (right panel). Panel A. includes results for all elections and Panel B. includes the same estimates just for MU firms. The p-values below the labels are from testing whether the effects for the first and third terciles are equal. Across both figures, the effect of unionization on overall employment growth and survival is larger for elections in the top tercile of the delay time distribution. Focusing on elections at MUs, the effects are significantly different for both outcomes at the three- and ten-year horizon. At the ten-year horizon, the estimated effect on survival for the top tercile is -0.20 (std. err. .037) versus -.071 (std. err. .036) for the bottom tercile. Finally, it is reassuring that conditioning on delay time, a proxy for the strength of the anti-union campaign, does not substantially reduce our overall estimates of the impacts of unionization.

We next check whether the delay time heterogeneity holds using a continuous measure of delay time. While the within-year terciles are appealing because they only rely on within-year variation in delay time and allow for a flexible, functional form, we might have more power using the entire distribution of delay times. Specifically, we add to specification equation 8 an interaction between the event-time treatment indicators times the log filing time for each election. Table 6 presents the coefficient estimates on the log election delay interaction for three, five, and ten years post-election. The first two columns show that the negative effects of unionization are significantly larger for elections with longer delays across all time horizons. For the ten-year survival effect, an approximately 10% increase in election delay is associated with a .7 pct. pt. increase in the probability of a plant closing. Columns (3) and (4) show that the effects are robust to including the controls pooled across cohorts. Columns (5) and (6) address the concern that our election delay time measure is just capturing larger bargaining units. Specifically, the estimates show that estimates are qualitatively, although no longer significant at the ten-year horizon, the same when we first residualize the log delay time on bargaining unit size deciles.

Our primary interpretation of these results is that the negative effects of unionization are largest at the establishments where the employer initially campaigned harder against the union. This is supported by anecdotal accounts linking election delays to the intensity of firms’ anti-union campaigns (Levitt and Conrow, 1993). Another interpretation of these results is that delay may be a proxy for hostile labor relations conditions (e.g., conflict between the union and firm before the election could delay it). Overall, this heterogeneity adds to our results suggesting that managers’

---

57 We also control for log filing time interacted with event-time directly in the specification.
58 Reassuringly, the magnitudes of the continuous-specification estimates are similar to the tercile specification estimates. The implied survival difference from the continuous specification between the 10th and 90th percentiles of the delay time difference is $[\ln(80) - \ln(30)] \times -0.07 = -0.066$. The ten-year survival difference between the first and third terciles is $-0.089$.
59 A different explanation of the delay time is that union organizer mistakes could lead to longer election delays (e.g., filing a petition for an improper bargaining unit). Yet, it is not clear why this would also lead to larger negative treatment effects and this explanation is not featured in the many analyses of the influence of delay time on election outcomes.
opposition to unions plays a role in the overall negative effects of successful union elections.

**Unionization and Productivity Reallocation** Finally, we examine how the negative effects of unionization vary by establishment productivity. The motivation for this analysis is that theoretical and empirical work in other contexts predicts that wage increases or productivity declines should have larger impacts at lower productivity establishments. Since these channels are two leading “economic” reasons why unionization might cause decreased employment or exit, this work suggests that unionization may also have a larger impact on lower productivity plants. Consequently, substantial heterogeneity by establishment productivity may be more consistent with the survival effects being driven by direct effects on wages or productivity than our union avoidance hypothesis.

To measure establishment-specific TFP for our manufacturing elections, we use cost-share-based productivity measures from the Annual Survey of Manufacturers (ASM) calculated by Foster et al. (2016). We use within-industry TFP comparisons to address potential measurement or productivity differences across industries. Specifically, we classify each election establishment into three productivity terciles based on their pre-election, within year and six-digit NAICS industry TFP ranking (see Appendix Section C for details). Figure 11 plots the estimated effects for the first and third terciles of the baseline TFP distribution. We find evidence that the three- and five-year employment and exit effects are larger for lower-productivity establishments. But these differences are never significant and at the five-year and ten-year horizon are not economically very large (e.g., -.066 (std. err. .023) versus -.041 (std. err. .022) at the five-year horizon for the first and third terciles, respectively). Appendix Figure A6 shows that these patterns hold separately estimating heterogeneity by baseline TFP only for MU establishments (e.g., larger but insignificant three-year effects for less productive establishments that do not persist). Overall, we interpret this evidence as only suggestive of larger short-run exit effects for less productive establishments. Thus, the evidence is more consistent with alternative explanations for why unionization leads to plant closures (e.g., our union avoidance hypothesis) than conventional explanations.

7 Relation to Literature and Implications

In this section, we relate our overall estimates to the literature on unionization’s employment and survival effects. Additionally, we discuss the implications of our union avoidance results for interpreting these effects.

---

60For the effect of wage increases, Berger et al. (2021) show that minimum wage increases can cause relatively larger employment declines at less productive firms in oligopsonistic labor markets with firm heterogeneity. Dustmann et al. (2020) confirm this by finding that a minimum wage increase in Germany caused firm exit and shifted employment towards larger and higher-productivity firms. We note, however, that the settings are different as minimum wages are market-wide wage increases while unionization would be an establishment-specific wage increase. We are not, however, aware of any models of unionization with firm heterogeneity and imperfectly competitive labor markets. For the effect of productivity declines, see (Foster et al., 2008) for evidence that plants’ level of revenue-based productivity is a key determinant of exit. One additional caveat to extending both predictions to the effects of unionization is that unions could base their bargaining demands on establishment productivity (e.g., try to extract more from higher productivity establishments).
Unionization, Employment, and Survival Literature. Our results add to a recent literature finding that unionization over the past four decades has led to establishment-level employment and survival likelihood declines. Specifically, we find that five years after an election, winning establishments are four pct. pts. less likely to survive than similar losing establishments. The most comparable results to ours are Frandsen (2021)’s regression discontinuity estimates also using the LBD. We qualitatively match his short-run employment and long-run survival declines but find somewhat smaller effects (e.g., five-year survival effects of 4 pct. pts. versus 8-10 pct. pts.). Some explanations for this are different samples or different empirical strategies and identifying assumptions. Additionally, the pattern we find of small employment declines but insignificant survival effects for the service sector is the same as Sojourner et al. (2015)’s estimates of negative employment but no survival effects following nursing home elections. However, our estimates, even for relatively close elections, are inconsistent with DiNardo and Lee (2004)’s null effects on both establishment survival and employment.

Implications of Firm Union Avoidance and Opposition. For manufacturing, we present evidence suggesting that firms’ union avoidance tactics and their opposition to unions play a role in these overall negative effects. First, we find evidence consistent with firms avoiding new unions by shifting production away from newly unionized plants. Second, we find larger effects at firms more opposed to unions based on two different proxies (non-unionized firms and long delay elections).

One interpretation of these results is that firms’ opposition to unions and attempts to avoid unions are completely driven by managers’ dislike of working with unions and unrelated to any direct costs of unions. This interpretation is consistent with accounts suggesting that our measures of manager opposition were motivated by manager dislike. For example, Freedman (1979) finds that non-unionized firms placed the most weight on resisting new unions but that these firms were also where unions were least able to attain higher wages (e.g., the threat of strikes was more limited because these firms could still produce at a non-union plant). Similarly, Foulkes (1980) documents that some non-unionized firms were motivated by a philosophical opposition to unions even if they did not have previous bad experiences with unions. Finally, Bronfenbrenner (2001) finds that the intensity of firms’ anti-union campaigns was “unrelated to the financial condition of the employer, but rather were a function of the extreme atmosphere of anti-union animus.” This interpretation also helps resolve the puzzle that it has been difficult to find evidence of unionization raising wages or negatively affecting productivity but that we find large effects on establishment survival. Additionally, it is puzzling that unions would not reduce their bargaining demands to avoid these effects (Friedman, 1951). The manager dislike interpretation provides a resolution to

61 For example, Frandsen (2021) requires at least 20 votes cast to implement the regression discontinuity analysis while we only require more than five eligible voters. This could explain the differences in the magnitudes of our estimates since the average election in our sample may have a lower bargaining unit to overall establishment employment share and we find larger treatment effects for elections where the bargaining unit is a larger share of overall employment.

62 One explanation is that the longitudinal linkages necessary to define survival in the LBD may be much higher quality than in the telephone book based InfoUSA or the LRD data used in DiNardo and Lee (2004). See, Jarmin and Miranda (2002) and Crane and Decker (2019) for comparisons between these datasets and descriptions of the quality of the longitudinal linkages.
these puzzles as the survival effects may not be driven by direct costs of unions.

An intermediate interpretation of our results is that while there may be some direct costs unionization, the survival and employment effects we estimate are excessive relative to these costs. For example, if firms can easily shift production away from unionized establishments, even small wage or productivity effects could lead to large survival effects. Additionally, the larger effects we find at non-unionized firms could be due to efforts to prevent unionization from spreading to other establishments by establishing a reputation of being hard on unions.\textsuperscript{63} One channel that could magnify such effects is evidence of unionization directly affecting managers’ careers. For example, (Clark, 1980) finds increased manager turnover after successful elections and (Dunlop, 1994) document some managers’ expectations that unionization would hurt their career prospects.

On the other hand, we cannot rule out that our proxies for manager opposition simply reflect rational expectations of where unions would have been the costliest. This interpretation is supported by evidence suggesting direct costs of unions (e.g., stock price declines following successful union elections and a large body of literature on unions reducing firm profits (Freeman and Medoff, 1984; Lee and Mas, 2012)). However, even parts of Lee and Mas (2012)’s stock price results are difficult to reconcile with our results. In particular, as noted in Frandsen (2021), Lee and Mas (2012) only find equity value declines away from the 50\% threshold while we and Frandsen (2021) also find survival and employment declines for close elections. A potential reconciliation is that the survival effects for these very close elections may not be driven by direct costs.

Overall, since we do not provide estimates of the magnitude of the direct costs of unions, we cannot disentangle these different interpretations. Yet, our evidence that the largest negative effects of unionization are where firms are the most opposed to unions and where they can potentially avoid unions, suggests that the overall negative effects may not necessarily imply large direct costs.

8 Conclusion

In this paper, we revisit the effects of successful NLRB union elections on establishments’ employment and survival. We first show that winning union recognition decreases establishments’ employment and long-run survival. While one interpretation of these results is that unions must have large direct costs on businesses, we raise alternative explanations for the effects and explore whether our results for manufacturing elections are consistent with these explanations. First, the overall negative effects may be partially driven by managers’ dislike of working with unions. Supporting this, we find that the largest effects at non-unionized firms and at elections with the longest pre-election delay – both proxies for a firms’ opposition to unions. Second, we hypothesize that the effects of unionization could be magnified by firms’ ability to avoid working with the union. We support this hypothesis by showing that the largest effects are at multi-establishment firms and by providing direct evidence of spillover effects on firms’ other plants. Both results are consistent with firms shifting employment away from newly unionized plants, one version of avoiding working with unions. Overall, our

\textsuperscript{63}As an extreme example, consider Walmart’s switch to pre-packaged meat across all stores days after ten Walmart meat cutters at one Texas store voted to unionize in 2000 (Zimmerman, 2000).
findings indicate that understanding how firms can respond to unionization and managers’ reasons for opposing unionization are important for interpreting the overall impact of unions.

Finally, our results raise many questions about the impact of unionization that suggest opportunities for further research. First, our results highlight how firms that strongly oppose unions go to great lengths to avoid working with unions. Another strategy they may pursue is raising workers’ wages to discourage unionization (i.e., “union threat effects”) which may be one of the main channels through which unions have affected the U.S. wage structure (Taschereau-Dumouchel, 2020). Further research quantifying this channel would be helpful to understand unions’ total effect on wages including these indirect effects (Neumark and Wachter, 1995; Farber, 2005). Second, we present evidence of large differences in how firms in different sectors respond to unionization. These findings mirror the heterogeneity in how manufacturing versus service-sector firms respond to minimum wage increases (Cengiz et al., 2019; Harasztosi and Lindner, 2018). These results suggest that understanding how more tradable industries respond to wage increases would be useful for understanding the effects of mandating higher wages in these industries. Specifically, the production shifting channel supported in this paper may also help explain why other labor market policies have heterogeneous effects across sectors.
References


Bronfenbrenner, Kate (1991) “Successful Union Strategies For Winning Certification Elections And First Contracts: Report To Union Participants (Part 2: First Contract Survey Results).”


Figure 1: Testable Implications of Parallel Trends Identifying Assumption

Note: This figure illustrates our empirical strategy’s identifying assumption and its testable implications discussed in section 4. It plots hypothetical average establishment-level outcomes before and after union elections with different vote shares. $Y_{i,-2}$ and $Y_{i,-1}$ correspond to outcomes one and two years before the union election. $Y_{i,1}$ corresponds to outcomes one year after the election. Testing parallel pre-trends by vote share corresponds to comparing the distance between $Y_{i,-2}$ and $Y_{i,-1}$. Testing parallel post-trends for losing elections corresponds to comparing the distance between $Y_{i,-2}$ and $Y_{i,1}$ for losing elections.

Figure 2 Note: Figure 2 present four panels illustrating characteristics of close union elections. All panels are constructed using external union election data (e.g., not our final sample matched to the Census) but the sample was constructed to mirror the overall sample construction (see Appendix Section C for details). Panel A. plots the vote-share histogram of elections included in our sample. Given the discreteness of the running variable and the fact that our sample includes elections with a small number of votes, it is difficult to detect manipulation from the vote-share density figure. Consequently, we plot elections with exactly 50% of votes separately to make the manipulation easier to see. See Frandsen (2017) for evidence of manipulation using formal tests that accommodate discrete running variables. Panel B. plots the average and median number of days between the union election date and the date that the case closed. Panel C. plots the probability of each union election experiences a decertification election in the five years following the case closing. The decertification elections are also from our combined NRLB datasets but excluded from our main analysis. Panel D. plots the probability of each union election experiencing a works stoppage in the five years following the case closing. The works stoppage data is from the FMCS and covers works stoppages from 1984-2019. Consequently, we only plot follow-up works stoppages for elections from 1984-2005. For the decertification and works stoppage figures, we match based on exact company names and cities rather than the SoftTFIDF algorithm we use for the main analysis. The “conditional regression coefficients” are the coefficients from regressing the stoppage indicators on the vote share for winning elections including controls for deciles of the number of workers in the bargaining unit, the four-digit NAICS industry, and election state.
Figure 2: Characteristics of Close Elections that Motivate Including Larger Margin of Support Elections

Panel A. Election Vote Share Histogram
Density of Union Certification Elections

Panel B. Number of Days Between Election and Case Closing Dates
Days Between Election Date and Case Closing Date

Panel C. Probability Decertification Election Five Years Following Election
Share of Elections with Decertification Attempt within 5 Years

Panel D. Probability of Works Stoppage Five Years Following Election
Share with Works Stoppage 5 Yrs Post Election

Note: See the previous page.
Figure 3: Employment and Survival Estimates, 20-80 % Vote Shares, All Industries

Panel A. DHS Employment Growth

Panel B. Log Employment

Panel C. Employment, Payroll, and Survival Estimates

Note: This figure plots the $\delta_n$ coefficients (i.e., the interaction between winning a union election and being $n$ years from the election) from estimating specification 8 for all union elections with 20-80 % vote shares inclusive. The sample includes observations -10 to 10 years before and after each union election but we only plot the -5 to 5 coefficients. The outcome variable for Panel A. is establishment-level DHS employment growth relative to time $-1$. The outcome variable for Panel B. is establishment-level log employment. The outcome variables for Panel C. are DHS employment and payroll growth rates and an indicator for whether the establishment exists at time $t$. For Panel C., the survival y-axis is scaled to be one-half the DHS growth rate axis. Consequently, comparing the exit and DHS coefficients illustrates how much of the effect on the DHS growth rate can be mechanically explained by the exit effect. Panels A. and B. include estimates with no controls, just industry and employment controls, and the flexible control specification (see Section 4 for details). Panel C. includes estimates from the flexible control specification. The log outcome estimates in Panel B. include establishment fixed effects but these are not included in Panel A. or Panel C. Standard errors are clustered by establishments’ firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment).
Figure 4: Employment and Survival Estimates, 20-80 % Vote Shares, Manufacturing

Panel A. DHS Growth Rates

Panel B. Log Employment and Payroll

Panel C. Employment, Payroll, and Survival Estimates

Note: These estimates are identical to Figure 3 except that they are only estimated for manufacturing elections.
Figure 5: Nonparametric Vote Share Heterogeneity Estimates, Manufacturing

Panel B. DHS Employment Growth Rate

Note: This figure plots the $\delta_{g,n}$ coefficients from estimating the vote-share heterogeneity specification 10 with the vote-share distribution partitioned into eight groups indicated on the x-axis. We omit the 20-30 % election group so the other estimates are relative to that group. The sample includes all manufacturing elections. We include observations -10 to 10 years before and after each union election but we only plot a subset of coefficients. The outcome variable for Panel A. is establishment-level DHS employment growth relative to event time $-1$. The outcome variable for Panel B. is an indicator for establishment survival. The estimates include the flexible control specification (see Section 4 for details). Standard errors are clustered by establishments’ firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment).
Figure 6: Nonparametric Vote Share Heterogeneity Estimates, All Industries

Panel A. DHS Employment Growth Rate

\[
\begin{align*}
\text{DHS Employment Growth Rate} & = \begin{cases} 
\text{5 Yr} & = -0.035 (0.024) \\
\text{10 Yr} & = -0.051 (0.027)
\end{cases} \\
\text{Excluding 50 \% Votes:} & \begin{cases} 
\text{5 Yr} & = -0.015 (0.025) \\
\text{10 Yr} & = -0.032 (0.028)
\end{cases}
\end{align*}
\]

\[\text{Panel A. Establishment Survival}\]

Note: This figure is identical to Figure 5 except it includes elections across all industries. The alternative estimates listed in the text box in Panel A. are the 40-50 \% estimates excluding elections with exactly 50 \% of votes (rather than restrict the sample, we include a separate category for 50 \% vote elections).
Figure 7: Single- Versus Multi-Establishment Firm Heterogeneity

Note: This figure plots the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 12 for elections at single- versus multi-establishment firms. An election at a multi-establishment firm is defined based on whether the establishment’s firm has any other establishments one year before the election. The sample includes all manufacturing union elections with 20-80 % vote shares inclusive. It includes observations -10 to 10 years before and after each union election but we only plot a subset of these coefficients. The outcome variable for the left panel is DHS employment growth rates relative to time $-1$ (see Section 4 for their definition). The outcome variable for the right panel is an indicator for establishment survival. The estimates include the flexible control specification (see Section 4 for details). The control coefficients are pooled across the heterogeneity groups. See Appendix Table A5 for robustness to alternative controls specifications. Standard errors are clustered by establishments’ firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment).
Figure 8: Employment Effects of Successful Elections on Firms’ Other Establishments

Panel A. All Plants and Three-Digit NAICS Plants

Note: This figure plots the $\delta_n$ coefficients from estimating specification 8. The sample is manufacturing plants at multi-establishment firms where another plant experienced a union election. See Section 6 for details about the sample construction. The relative time and vote share variables are defined from the election at the firm’s other establishment. We weight the regression by the observation’s share of total firm-level employment across all plants included in the sample the year of the election. The outcomes in both panels are establishment-level DHS employment growth rates relative to one year before the union election. The estimates include the flexible control specification (see Section 4 for details) except we do not include a control for establishments SU/MU status (all plants are part of MUs) or for establishments’ previous contract status. Since we match plants based on the election year, the industry is also from the year of election. The “All Manufacturing Estabs” estimates in the left panel include all manufacturing establishments with at least two employees during the year of the election. The “Within-NAICS 3 Estabs” estimates restrict the sample to plants that are in the same 3-digit NAICS industry as the election plant. The right panel includes 3-digit NAICS industry matches but separately estimates the effects by whether or not the election establishment comprised more than 10 % of the firm’s employment in the same three-digit NAICS industry during the year of election. The estimates in Panel B. are from the same specification with the controls pooled across both groups and the treatment indicators interacted with the two employment share groups. In this panel, we also directly control for the effect of the two employment share groups interacted with event time.

Panel B. Estimates by Election’s Employment Share
Figure 9: Unionized versus Non-Unionized Firm Heterogeneity

Note: This figure plots similar heterogeneity results as Figure 7 except that the heterogeneity is for elections at multi-establishment firms with at least one unionized establishment versus firms without any unionized establishments. See Appendix Section C for how we define firms’ unionization status. The controls additionally directly include these heterogeneity groups interacted with cohort and event time.
Figure 10: Election Delay Heterogeneity

Panel A. All Elections

![Chart showing election delay heterogeneity for all elections.](chart)

- **DHS Employment Growth Rate**
  - First Delay Time Tercile: Red bars
  - Third Delay Time Tercile: Blue bars
  - 5-Yr Pre: p-val 0.038
  - 3-Yr Post: p-val 0.556
  - 5-Yr Post: p-val 0.092
  - 10-Yr Post: p-val 0.038

- **Establishment Survival**
  - 3-Yr Post: p-val 0.05
  - 5-Yr Post: p-val 0.23
  - 10-Yr Post: p-val 0.016

**Note:** These figures plots the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 12 for elections in different terciles of the election delay distribution. These terciles are defined within each year based on the number of days between the election petition filing date and the election date (see Section C for details). We plot the coefficients for the first and third terciles but estimate the effects for all three. The sample includes all manufacturing union elections with 20-80 % vote shares inclusive. It includes observations -10 to 10 years before and after each union election but we only plot a subset of these coefficients. The outcome variable for the left panel is DHS employment growth rates relative to time $-1$ (see Section 4 for their definition). The outcome variable for the right panel is an indicator for establishment survival. The estimates include the flexible control specification (see Section 4 for details). Standard errors are clustered by establishments’ firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment). Panel A. defines the election delay terciles across all elections. For Panel B. the election delay terciles are only defined for elections at multi-establishment manufacturing firms. Consequently, we estimate but do not report separate coefficients for elections at single-establishment firms.

Panel B. Elections at Multi-Establishment Firms

![Chart showing election delay heterogeneity for multi-establishment elections.](chart)

- **DHS Employment Growth Rate**
  - First Delay Time Tercile: Red bars
  - Third Delay Time Tercile: Blue bars
  - 5-Yr Pre: p-val 0.011
  - 3-Yr Post: p-val 0.134
  - 5-Yr Post: p-val 0.044

- **Establishment Survival**
  - 3-Yr Post: p-val 0.03
  - 5-Yr Post: p-val 0.211
  - 10-Yr Post: p-val 0.016

**Note:** These figures plots the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 12 for elections in different terciles of the election delay distribution. These terciles are defined within each year based on the number of days between the election petition filing date and the election date (see Section C for details). We plot the coefficients for the first and third terciles but estimate the effects for all three. The sample includes all manufacturing union elections with 20-80 % vote shares inclusive. It includes observations -10 to 10 years before and after each union election but we only plot a subset of these coefficients. The outcome variable for the left panel is DHS employment growth rates relative to time $-1$ (see Section 4 for their definition). The outcome variable for the right panel is an indicator for establishment survival. The estimates include the flexible control specification (see Section 4 for details). Standard errors are clustered by establishments’ firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment). Panel A. defines the election delay terciles across all elections. For Panel B. the election delay terciles are only defined for elections at multi-establishment manufacturing firms. Consequently, we estimate but do not report separate coefficients for elections at single-establishment firms.
Figure 11: Plant-Level Total Factor Productivity Heterogeneity

Note: This figure plots the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 12 for elections in different terciles of baseline TFP distribution. These terciles are defined based on plants’ pre-election cost-share-based productivity measures from the Annual Survey of Manufacturers (ASM) calculated by Foster et al. (2016). The TFP terciles are defined based on within-year and within six-digit NAICS productivity rankings. See Appendix Section C for more details. We plot the coefficients for the first and third terciles but estimate effects for all three terciles and a fourth group of plants without TFP defined. The sample includes all manufacturing union elections with 20-80 % vote shares inclusive. It includes observations -10 to 10 years before and after each union election but we only plot a subset of these coefficients. The outcome variable for the left panel is DHS employment growth rates relative to time $-1$ (see Section 4 for their definition). The outcome variable for the right panel is an indicator for establishment survival. The estimates include the flexible control specification (see Section 4 for details). The controls additionally include these heterogeneity groups interacted with cohort and event time. Standard errors are clustered by establishments’ firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment).
Table 1: Winning versus Losing Election Establishment Summary Statistics

<table>
<thead>
<tr>
<th>Establishment Characteristics</th>
<th>All Industries</th>
<th>Manufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Union Loses</td>
<td>Union Wins</td>
</tr>
<tr>
<td>Employees</td>
<td>154</td>
<td>137</td>
</tr>
<tr>
<td>Payroll/Worker ($ 2019)</td>
<td>49,400</td>
<td>49,700</td>
</tr>
<tr>
<td>Establishment Age</td>
<td>9.65</td>
<td>10.0</td>
</tr>
<tr>
<td>Multi-Establishment Firm</td>
<td>0.512</td>
<td>0.476</td>
</tr>
<tr>
<td>Previous Contract at Establishment</td>
<td>0.090</td>
<td>0.147</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Survival Base Rates</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>5-Year Survival</td>
<td>0.818</td>
<td>0.765</td>
</tr>
<tr>
<td>10-Year Survival</td>
<td>0.667</td>
<td>0.610</td>
</tr>
</tbody>
</table>

| Approximate Number of Elections | 27,000 | 7,000 |

*Note:* This table presents summary statistics for all union elections included in our analysis sample with vote shares between 0-100%. All establishment characteristics are measured one year before the union election. Since the FMCS contract data are only available starting in 1984, we only calculate the share of establishments with a previous contract using elections from 1985 onward. The five- and ten-year survival rates are the probability of surviving five and ten years after the union election respectively. To satisfy the Census’ disclosure requirements, all estimates are rounded to only include three significant digits and sample sizes are rounded to the nearest 1,000.
Table 2: Pre-Election Employment Growth Trends by Vote Share, 20-80 % Vote Shares

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>DHS Employment Growth Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industry Group:</td>
<td></td>
</tr>
<tr>
<td>5-Year Pre Election × Vote Share</td>
<td>All Industries</td>
</tr>
<tr>
<td></td>
<td>0.050</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
</tr>
<tr>
<td>4-Year Pre Election × Vote Share</td>
<td>0.018</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
</tr>
<tr>
<td>3-Year Pre Election × Vote Share</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
</tr>
<tr>
<td>2-Year Pre Election × Vote Share</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
</tr>
<tr>
<td>Industry + Employment Ctrls.</td>
<td>X</td>
</tr>
<tr>
<td>Flexible Ctrls.</td>
<td>X</td>
</tr>
<tr>
<td>Number of Elections</td>
<td>19,000</td>
</tr>
</tbody>
</table>

Note: This table presents estimates testing for linear trends by vote share in pre-election employment growth rates. Significant estimates would violate a testable implication of our parallel trends by vote share assumption (see equation 5). Specifically, the table reports the estimated coefficients on interactions between event-time indicators and the continuous election vote-share (i.e., the $\rho$ coefficients from Equation 11). A five-year coefficient of .03 implies that elections with 75 % of votes grew approximately 1.5 percent slower during the five years before the election than an election with 25 % of votes. The outcome for all specifications is establishment-level DHS employment growth relative to time $-1$. The sample includes 20-80 % vote-share elections. The first two columns include elections in all industries and the last two columns include just manufacturing elections. The odd columns include only industry and employment controls and the even columns include our flexible control specification (see Section 4 for details). Standard errors are clustered by establishments’ firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. 
Table 3: Post-Election Outcome Trends by Vote Share, 20-80 % Vote Shares

<table>
<thead>
<tr>
<th>Outcome</th>
<th>DHS Emp Growth Rate</th>
<th>Establishment Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Event-Time × 0-50 % Vote Share</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-Year Post Election</td>
<td>-0.216**</td>
</tr>
<tr>
<td></td>
<td>(0.095)</td>
<td>(0.103)</td>
</tr>
<tr>
<td></td>
<td>5-Year Post Election</td>
<td>-0.220**</td>
</tr>
<tr>
<td></td>
<td>(0.110)</td>
<td>(0.122)</td>
</tr>
<tr>
<td></td>
<td>10-Year Post Election</td>
<td>-0.332***</td>
</tr>
<tr>
<td></td>
<td>(0.125)</td>
<td>(0.140)</td>
</tr>
<tr>
<td></td>
<td>Event-Time × 50-100 % Vote Share</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-Year Post Election</td>
<td>-0.280**</td>
</tr>
<tr>
<td></td>
<td>(0.131)</td>
<td>(0.131)</td>
</tr>
<tr>
<td></td>
<td>5-Year Post Election</td>
<td>-0.381**</td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td>(0.149)</td>
</tr>
<tr>
<td></td>
<td>10-Year Post Election</td>
<td>-0.271*</td>
</tr>
<tr>
<td></td>
<td>(0.164)</td>
<td>(0.164)</td>
</tr>
</tbody>
</table>

Panel A: All Industries

| Event-Time × 0-50 % Vote Share |         |         |
| 3-Year Post Election | -0.236 | -0.145 | 0.017 | 0.035 |
| (0.159) | (0.170) | (0.061) | (0.065) |
| 5-Year Post Election | -0.216 | -0.072 | -0.023 | 0.025 |
| (0.187) | (0.199) | (0.076) | (0.081) |
| 10-Year Post Election | -0.425* | -0.210 | -0.151 | -0.049 |
| (0.226) | (0.241) | (0.097) | (0.104) |

Panel B: Manufacturing

| Event-Time × 0-50 % Vote Share |         |         |
| 3-Year Post Election | -0.462* | -0.470* | -0.009 | -0.010 |
| (0.266) | (0.266) | (0.104) | (0.104) |
| 5-Year Post Election | -0.394 | -0.406 | 0.008 | 0.004 |
| (0.299) | (0.299) | (0.126) | (0.126) |
| 10-Year Post Election | -0.559* | -0.578* | -0.162 | -0.171 |
| (0.336) | (0.336) | (0.150) | (0.150) |

Exclude 50 % Elections | X | X |
| Industry + EmploymentCtrls. | X | X | X | X |
| FlexibleCtrls. | X | X | X | X |

Note: This table presents estimates testing for linear trends by vote share in post-election outcomes. We test for trends separately across winning versus losing elections. The Event-Time × 0-50 rows present estimates of the θ coefficients from equation 11 and capture linear trends in post-election outcomes for losing elections. The Event-Time × 50-100 rows present estimates of θ + τ and capture linear trends in post-election outcomes for winning elections. Since the specification separately includes an interaction with a winning election indicator, these slope estimates are in excess of any treatment effect right around the 50 % threshold. The outcome for the first two columns is establishment-level DHS employment growth relative to time −1. The outcome for the last two columns is an indicator of whether the establishment exists at time t. All specifications include our flexible control specification (see Section 4 for details). See Appendix Table A2 for the same results with alternative included controls. The columns that “Exclude 50 % Elections” include an interaction between having a vote share of exactly 50 % and event time. * p < 0.1, ** p < 0.05, *** p < 0.01.
<table>
<thead>
<tr>
<th>Industry Group:</th>
<th>Manufacturing</th>
<th>Services</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DHS Emp</td>
<td>Survival</td>
<td>DHS Emp</td>
</tr>
<tr>
<td>5-Year Pre Election</td>
<td>0.005 (0.015)</td>
<td>0.010 (0.012)</td>
<td>0.011 (0.016)</td>
</tr>
<tr>
<td>2-Year Pre Election</td>
<td>-0.013 (0.012)</td>
<td>0.017* (0.009)</td>
<td>-0.009 (0.012)</td>
</tr>
<tr>
<td>5-Year Post Election</td>
<td>-0.174*** (0.029)</td>
<td>-0.047*** (0.012)</td>
<td>-0.057** (0.024)</td>
</tr>
<tr>
<td>10-Year Post Election</td>
<td>-0.231*** (0.033)</td>
<td>-0.075*** (0.015)</td>
<td>-0.059** (0.027)</td>
</tr>
<tr>
<td>Industry + Employment Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Flexible Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Industry Group Number of Elections</td>
<td>6,000</td>
<td>6,000</td>
<td>8,000</td>
</tr>
<tr>
<td>Industry Group Share of Elections</td>
<td>0.302</td>
<td>0.302</td>
<td>0.414</td>
</tr>
</tbody>
</table>

Note: This figure plots the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 12 for elections in three different broad industry groups. Manufacturing is defined as NAICS sectors 31-33, services are defined as NAICS 51-81 and retail trade (NAICS 44-45), and other is the remaining industries. Elections are classified into industries based on their Fort and Klimek (2016) NAICS 2012 codes. Otherwise, the sample, controls, and standard errors are the same as in Figure 3. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. 

52
Table 5: Effects of Successful Elections on Firms’ Other Establishments

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>DHS Employment</th>
<th>Survivial</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Year Post Election</td>
<td>0.017</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>2-Year Post Election</td>
<td>0.043**</td>
<td>0.012</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>3-Year Post Election</td>
<td>0.044**</td>
<td>0.022**</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>4-Year Post Election</td>
<td>0.048*</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>5-Year Post Election</td>
<td>0.034</td>
<td>0.023**</td>
</tr>
<tr>
<td></td>
<td>(0.025)</td>
<td>(0.011)</td>
</tr>
</tbody>
</table>

Industry + Employment Ctrls. X X
Flexible Ctrls. X X

Note: This table presents the DHS employment growth rate and survival estimates estimated as described for Figure 8. The DHS employment growth rate estimates exactly match the DHS employment estimates presented in that table.

Table 6: Election Delay Heterogeneity, Continuous Delay Time Specification

<table>
<thead>
<tr>
<th>Treatment:</th>
<th>Log Delay Time</th>
<th>Residualized Log Delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome:</td>
<td>DHS Emp</td>
<td>Survival</td>
</tr>
<tr>
<td>3-Year Post Election</td>
<td>-0.124**</td>
<td>-0.057**</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>5-Year Post Election</td>
<td>-0.121*</td>
<td>-0.064**</td>
</tr>
<tr>
<td></td>
<td>(0.063)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>10-Year Post Election</td>
<td>-0.147**</td>
<td>-0.071**</td>
</tr>
<tr>
<td></td>
<td>(0.073)</td>
<td>(0.033)</td>
</tr>
</tbody>
</table>

Industry + Employment Ctrls. X X X X X X
Pooled Ctrls. X X
Flexible Ctrls. X X

Observations: 6,000 6,000 6,000 6,000 6,000 6,000

Note: This table presents coefficient estimates from a modified version of specification 8. Specifically, we interact the treatment by event time indicators with the continuous log delay time. See Appendix Section C for details on how we calculate the delay time. The table reports the coefficients on these interactions at various time horizons. Thus, a survival coefficient of -0.05 means that the effect of successful unionization on survival is .5 pct. pts. higher for elections with a 10% longer delay time. The first four columns use the raw number of days between petition filing and election dates to define the log delay time. For the last two columns, we first regress log delay time on within-year deciles of the election bargaining unit size and use the residuals from this regression as the interaction. The sample includes all elections at manufacturing establishments -10 to 10 years before and after each union election but we only include a subset of these coefficients. The even columns include the DHS employment growth rate relative to time -1 as the outcome variable (see Section 4 for their definition). The odd columns include an indicator for whether the establishment exists at time t as the outcome. Standard errors are clustered by establishments’ firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment). * p < 0.1, ** p < 0.05, *** p < 0.01.
Appendix for

Unionization, Employer Opposition, and Establishment Closure
A Appendix Figures

Figure A1: Number of Unique Case Numbers Across Datasets versus NLRB Annual Reports

Note: This figure plots the total number of unique NLRB election cases each year in our data set and in the annual NLRB reports. These include all case types (e.g., ‘RC’ cases and non-RC cases) Our data set is from combining union election datasets from Henry Farber, J.P. Ferguson, and Thomas Holmes and publicly available data from the NLRB and picking one observation for each NLRB case number. See Appendix Section C for more details on our data construction process.
Note: This plots the vote-share histogram of elections with more than 50 total voters. Given the discreteness of the running variable and the fact that our sample includes elections with a small number of votes, it is difficult to detect manipulation from the vote-share density figure for the entire sample so we restrict the sample to elections with at least 50 votes. The figure was constructed using external union election data (e.g., not our final sample matched to the Census) but the sample was constructed to mirror the overall sample construction (see Appendix Section C for details). Note, there may still be a small bias from the “integer problem” described in DiNardo and Lee (2004) that could lead to an excess mass of elections right below 50% but simulations suggest that it is quite small with at least 50 votes.
Figure A3: Log Employment and Payroll Estimates, 20-80 % Vote Shares

Panel A. All Union Elections

Note: This figure plots estimates from the Flexible Controls specification presented in Figure 3 Panel B. and Figure 4 Panel B. The log employment estimates are identical to the estimates in Figures 3 and 4 but the log payroll estimates are not otherwise reported.
Figure A4: DHS Employment Estimates, 20-80 % Vote Shares, 10 Yr Pre- and Post-Periods

DHS Employment Growth Rate

Note: This figure plots the same DHS employment growth rate estimates as in Figure 3, Panel C and Figure 4, Panel C but includes the -10 to -5 pre-period estimates and the 6 to 10-year post-period estimates. Note, the panel is balanced from -5 years pre-election to 10 years post-election but not from -10 to -5 years pre-election. Consequently, each of the -5 to -10 point estimates average over slightly different cohorts.
Figure A5: Election Win Rates and Challenged Vote Rates by Delay Time

Panel A. All Elections

![Graph showing the relationship between pre-election delay times, election win rates, and challenged votes in elections. Pre-election delay times are defined as the number of days between the election petition being filed and the election date. The within-year percentiles of the election delay distribution are plotted on the x-axis. The share of elections with a challenged vote is defined as an indicator for any vote in the election being challenged. The sample of elections includes all elections in our “external elections dataset” described in Appendix Section C. The conditional regression coefficients are from regressing the election win indicator (or challenged vote indicator) on deciles of the number of eligible voters in the election, four-digit NAICS industry fixed effects, and election state fixed effects.]

Conditional Regression Coefficients x 100: Winning = -.088 (.007). Challenged = .067 (.007).

Panel B. Manufacturing

![Graph showing the relationship between pre-election delay times, election win rates, and challenged votes in elections. Pre-election delay times are defined as the number of days between the election petition being filed and the election date. The within-year percentiles of the election delay distribution are plotted on the x-axis. The share of elections with a challenged vote is defined as an indicator for any vote in the election being challenged. The sample of elections includes all elections in our “external elections dataset” described in Appendix Section C. The conditional regression coefficients are from regressing the election win indicator (or challenged vote indicator) on deciles of the number of eligible voters in the election, four-digit NAICS industry fixed effects, and election state fixed effects.]


Note: This figure plots the relationship between pre-election delay times, election win rates, and challenged votes in elections. Pre-election delay times are defined as the number of days between the election petition being filed and the election date. We then take the within-year percentiles of the election delay distribution and plot this on the x-axis. The share of elections with a challenged vote is defined as an indicator for any vote in the election being challenged. The sample of elections includes all elections in our “external elections dataset” described in Appendix Section C. The conditional regression coefficients are from regressing the election win indicator (or challenged vote indicator) on deciles of the number of eligible voters in the election, four-digit NAICS industry fixed effects, and election state fixed effects.
Note: This figure plots the same estimates as in Figure 11 except restricting the TFP comparison to only be between multi-establishment firms in different TFP terciles. As with the other heterogeneity tests, the sample includes all manufacturing elections and pools the controls across the entire sample.
## B Appendix Tables

### Table A1: Union Election Matched Sample Construction

<table>
<thead>
<tr>
<th>Panel A: NLRB Election Sample</th>
<th>All Elections</th>
<th>Winning Elections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Elections</td>
<td>Eligible Voters</td>
</tr>
<tr>
<td>All Election, 1981-2010</td>
<td>94,824</td>
<td>5,991,865</td>
</tr>
<tr>
<td>Representation Elections (RC)</td>
<td>77,349</td>
<td>5,111,675</td>
</tr>
<tr>
<td>&gt; 5 Eligible Voters</td>
<td>69,789</td>
<td>5,084,061</td>
</tr>
<tr>
<td>Non-Contested Elections</td>
<td>66,353</td>
<td>4,590,121</td>
</tr>
</tbody>
</table>

### Panel B: Final NLRB Sample Industry Shares

<table>
<thead>
<tr>
<th>Industry</th>
<th>All Elections</th>
<th>Winning Elections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.307</td>
<td>0.253</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>0.408</td>
<td>0.307</td>
</tr>
<tr>
<td>Other</td>
<td>0.266</td>
<td>0.263</td>
</tr>
<tr>
<td>Services</td>
<td>0.426</td>
<td>0.484</td>
</tr>
<tr>
<td></td>
<td>0.186</td>
<td>0.177</td>
</tr>
<tr>
<td></td>
<td>0.405</td>
<td>0.515</td>
</tr>
</tbody>
</table>

### Panel C: Matched Census Sample

<table>
<thead>
<tr>
<th>Sample</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elections Matched to Census Establishments</td>
<td>46,000</td>
</tr>
<tr>
<td>Final Establishment-Level Outcome Sample</td>
<td>27,000</td>
</tr>
<tr>
<td>20-80 % Election Sample</td>
<td>19,000</td>
</tr>
</tbody>
</table>

Note: This table illustrates how our specific sample restrictions change the number of elections and eligible voters we have in our sample. Panel A plots the total number of elections and eligible voters for all elections and specifically for winning elections. The first row in Panel A includes all unique NLRB cases with filing dates between 1981-2005 (the main years in our sample). The second row only includes representation (RC) elections. The third row drops elections without more than five eligible voters. The fourth row only includes non-contested elections (e.g., elections with one union on the ballot). Panel B presents the industry composition of the remaining elections from the fourth row of Panel A. Note we use the NLRB election industry codes here rather than the LBD industry codes but the overall industry shares are reassuringly similar to the industry shares in Table 4. The three columns represent the total shares of elections and eligible voters for all elections and winning elections. Panel C shows our final sample sizes from the matched Census data. The sample restrictions between “Elections Matched to Census Establishments” and “Final Establishment-Level Outcome Sample” include keeping (1) the first election at each establishment, (2) at least three years of pre-election survival, (3) non-missing employment, payroll, and other controls at event time $t = -1$. 

61
Table A2: Post-Election Outcome Trends by Vote Share, 20-80 % Vote Shares, Employment and Industry Ctrls.

<table>
<thead>
<tr>
<th>Industry Group:</th>
<th>All Industries</th>
<th>Manufacturing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome:</td>
<td>DHS Emp</td>
<td>Survival</td>
</tr>
<tr>
<td>3-Year Post Election</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event-Time × 0-50 % Vote Share</td>
<td>-0.134</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.037)</td>
</tr>
<tr>
<td>Event-Time × 50-100 % Vote Share</td>
<td>-0.361***</td>
<td>-0.052</td>
</tr>
<tr>
<td></td>
<td>(0.126)</td>
<td>(0.051)</td>
</tr>
<tr>
<td>5-Year Post Election</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event-Time × 0-50 % Vote Share</td>
<td>-0.119</td>
<td>-0.009</td>
</tr>
<tr>
<td></td>
<td>(0.116)</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Event-Time × 50-100 % Vote Share</td>
<td>-0.450***</td>
<td>-0.085</td>
</tr>
<tr>
<td></td>
<td>(0.141)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>10-Year Post Election</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event-Time × 0-50 % Vote Share</td>
<td>-0.218</td>
<td>-0.052</td>
</tr>
<tr>
<td></td>
<td>(0.133)</td>
<td>(0.057)</td>
</tr>
<tr>
<td>Event-Time × 50-100 % Vote Share</td>
<td>-0.354**</td>
<td>-0.107</td>
</tr>
<tr>
<td></td>
<td>(0.157)</td>
<td>(0.070)</td>
</tr>
<tr>
<td>Exclude 50 % Elections</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Industry + Employment Ctrls.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Flexible Ctrls.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of Elections</td>
<td>19,000</td>
<td>19,000</td>
</tr>
</tbody>
</table>

Note: This table presents the same estimates as in Tables 3 but only includes the baseline industry and employment controls. * p < 0.1, ** p < 0.05, *** p < 0.01.

Table A3: Employment and Survival Bargaining Unit Share Interaction, 20-80 % Vote Shares

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>DHS Employment</th>
<th>Survival</th>
</tr>
</thead>
<tbody>
<tr>
<td>3-Year Post Election × Bargaining Unit Share</td>
<td>-0.109**</td>
<td>-0.046***</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>5-Year Post Election × Bargaining Unit Share</td>
<td>-0.132***</td>
<td>-0.041*</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.021)</td>
</tr>
<tr>
<td>10-Year Post Election × Bargaining Unit Share</td>
<td>-0.057</td>
<td>-0.015</td>
</tr>
<tr>
<td></td>
<td>(0.057)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>Industry + Employment Ctrls.</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Flexible Ctrls.</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Note: This table presents estimates from the same specification as Figure 3 for DHS employment growth rates except that we add (1) an interaction between the event-time × win indicators with the share of the establishment’s employment covered by the bargaining unit and (2) an interaction just between event-time indicators and the bargaining unit share. We report the interactions in (1) for three, five, and ten years post election. Consequently, this specification estimates how treatment-effects increase with the bargaining unit share accounting for overall post-election trends across all elections by bargaining unit share. A survival estimate of -.05 means that increasing the share of the establishment covered by the bargaining unit by 10 % leads to an additional .5 pct. pct. increase in establishment exit.
Table A4: Manufacturing versus Services Employment and Survival Estimates, Robustness Checks

<table>
<thead>
<tr>
<th>Specification:</th>
<th>Baseline</th>
<th>Pooled Controls</th>
<th>Good Matches</th>
<th>&gt; 25 % Barg Unit Share</th>
<th>30-70 %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DHS Emp</td>
<td>Survival</td>
<td>DHS Emp</td>
<td>Survival</td>
<td>DHS Emp</td>
</tr>
<tr>
<td>5-Year Difference</td>
<td>-0.117</td>
<td>-0.021</td>
<td>-0.118</td>
<td>-0.022</td>
<td>-0.144</td>
</tr>
<tr>
<td></td>
<td>(0.037)</td>
<td>(0.016)</td>
<td>(0.035)</td>
<td>(0.015)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>10-Year Difference</td>
<td>-0.172</td>
<td>-0.058</td>
<td>-0.159</td>
<td>-0.05</td>
<td>-0.196</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.019)</td>
<td>(0.04)</td>
<td>(0.018)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>Industry + Employment Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Pooled Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Flexible Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Note: This table presents robustness results for the differences between the service-sector and manufacturing results in Table 4. Specifically, it presents the differences between the five- and ten-year DHS employment growth rate and survival estimates for various alternative specifications. The first two columns present the differences for the estimates presented in Table 4. The "Pooled Controls" columns pool the controls across all cohorts as described in Section 4. The "Good Matches" columns restrict to election matches which we give a 95 % rating (see Appendix Section C for details). The "Barg Unit Share" columns restrict to elections where the bargaining unit is at least 25 % of the total establishment employment. The 30-70 % columns restrict to elections with 30-70 % of the vote share. For all specifications with restrictions, we still use the entire sample for controls but restrict the treated variables to be estimated from the restricted sample.

Table A5: Single- Versus Multi-Establishment Firm Heterogeneity, Robustness Checks

<table>
<thead>
<tr>
<th>Specification:</th>
<th>Baseline</th>
<th>Pooled Controls</th>
<th>30-70 %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DHS Emp</td>
<td>Survival</td>
<td>DHS Emp</td>
</tr>
<tr>
<td>5-Year Difference</td>
<td>-0.068</td>
<td>-0.061</td>
<td>-0.063</td>
</tr>
<tr>
<td></td>
<td>(0.058)</td>
<td>(0.024)</td>
<td>(0.054)</td>
</tr>
<tr>
<td>10-Year Difference</td>
<td>-0.149</td>
<td>-0.093</td>
<td>-0.13</td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td>(0.03)</td>
<td>(0.062)</td>
</tr>
<tr>
<td>Industry + Employment Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Pooled Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Flexible Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Note: This table presents robustness results for the differences between single- and multi-establishment firms presented in Figure 7. Specifically, it presents the differences between the five- and ten-year DHS employment growth rate and survival estimates for various alternative specifications. The first two columns present the differences for the estimates presented in Figure 7. The "Pooled Controls" columns pool the controls across all cohorts as described in Section 4. The 30-70 % columns restrict to elections with 30-70 % of the vote share. For all specifications with restrictions, we still use the entire sample to estimate controls but restrict the treated variables to be estimated from the restricted sample.
Table A6: Unionized versus Non-Unionized Firm Heterogeneity, Robustness Checks

<table>
<thead>
<tr>
<th>Specification:</th>
<th>Baseline</th>
<th>Pooled Controls</th>
<th>Contracts since 1990</th>
<th>30-70 % Elections</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DHS Emp</td>
<td>Survival</td>
<td>DHS Emp</td>
<td>Survival</td>
</tr>
<tr>
<td>5-Year Difference</td>
<td>-0.187</td>
<td>-0.03</td>
<td>-0.139</td>
<td>-0.018</td>
</tr>
<tr>
<td></td>
<td>(0.095)</td>
<td>(0.041)</td>
<td>(0.089)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>10-Year Difference</td>
<td>-0.336</td>
<td>-0.108</td>
<td>-0.287</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td>(0.104)</td>
<td>(0.048)</td>
<td>(0.097)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>Industry + Employment Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Pooled Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Flexible Ctrls.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Note: This table presents robustness results for the differences between multi-establishment firms with and without any unionized establishments presented in Figure 9. Specifically, it presents the differences between the five- and ten-year DHS employment growth rate and survival estimates for various alternative specifications. The first two columns present the differences for the estimates presented in Figure 9. The "Pooled Controls" columns pool the controls across all cohorts as described in Section 4. The "Contracts since 1990" column only classifies firms as unionized versus non-unionized starting in 1990. This gives all firms at least five years of pre-election FMCS contract data that we can use to define the firms’ unionization status. The 30-70 % columns restrict to elections with 30-70 % of the vote share. For all specifications with restrictions, we still use the entire sample to estimate controls but restrict the treated variables to be estimated from the restricted sample.
C Data and Matching Details Appendix

Union Election Data We combine datasets on NLRB elections from Henry Farber, J.P. Ferguson, and Thomas Holmes and publicly available data from the NLRB to give us a near-complete set of union elections from 1961-2019. Internet links for the Ferguson, Holmes, and NLRB are available.

We deal with multiple observations per case number within datasets somewhat differently across datasets. For the public NLRB data (the “Public Data”), for a given bargaining unit, we pick the final tally of the last election for each case number. This ensures that in the case of recounts or multiple elections for the same bargaining unit that we take the vote tally that determines the unions’ eventual certification. Within each case number, we then take the results from the election at the largest bargaining unit in cases where there are distinct bargaining units for a single case. For the other datasets, there is somewhat less clarity about why there are duplicate observations within the same case number. For these datasets, we first pick the observation with the last election date and then the observation with the largest bargaining unit size.

This leaves us one observation per case number within each data set but duplicates across datasets. We take one observation per case number across datasets deprioritizing observations in the Farber data given data irregularities in those data and prioritizing the public data. We define the following from the union election data that we use for our analysis and for our matching algorithm

- **Election City, State, and Address:** for the public data, we observe these data both for the employer and for the election site. When they disagree, we choose the election site values as this is more likely to match the establishment unionizing rather than a separate corporate headquarters.

- **Election Filing Date:** We define treatment based on the date that the election is filed. For 5% of elections, we do not observe the filing date and instead use the election or case closing date.

- **Election Delay Time:** We define delay time as the number of days between the date the election petition was filed to the NLRB and the date the election was held. The availability of exact dates for these two concepts varies somewhat across time and datasets. Both dates are missing from the Farber data which is one reason why we prioritize the other datasets when duplicates across case numbers are available. For the Ferguson and Holmes data, the delay time is missing for cases that closed in 1982 and we only have a monthly measure for 1981 and part of 1983. These differences over time motivate our checks that the heterogeneity by delay time holds using both variation within-years (e.g., the within-year tercile measures) and across years (e.g., the continuous log specification). Additionally, there may have been some institutional changes over time that we do not want to include (e.g., the “Quickie Election Rule” decreased delay times but is not in our sample of elections).

- **Bargaining Unit Share:** We define the bargaining unit share as the number of eligible voters from the NLRB election data divided by establishment employment one year before the
election. Since we do not impose that the bargaining unit is smaller than the establishment, we cap the share at one.

FMCS Contract Data: We combine contract data from Thomas Holmes for 1984-2003 and from the FMCS for 1997-2019. They include both notices of *initial contracts* (i.e., first-contract negotiation after an election) and *contract renegotiation or reopening* for existing contracts. These “notices of bargaining” are provided to the FMCS so it can be ready to provide mediation. Although filing is legally incentivized, underreporting is possible. For example, an employer changing the terms of employment or a union striking without first filing a notice could be violating labor law.

There are duplicate observations both across datasets and within datasets (e.g., an employer and union might both submit an FMCS notice for the same contract). However, unlike the NLRB election data, we have no IDs to restrict the dataset to unique observations. Consequently, we match all contract observations to the LBD and drop duplicates when multiple contract observations match to the same LBDNUM.

We use the contract data to define

- **Previous contract at own establishment**: indicator for whether we ever matched an FMCS contract to this LBDNUM.

- **Unionized versus Nonunionized Firms**: we define a firm as being (partially) unionized if at time $t$ any of the establishments in the same FIRMID had an FMCS contract match in the current or previous five years. When defining this measure for the heterogeneity check, we exclude contracts matched to the election establishment during the year of the election to avoid contracts resulting from the election.

Longitudinal Business Database In Section 3, we mention potential concerns with how the LBD allocates employment across establishments at multi-establishment firms. To be more precise about the issue, one source used to construct the LBD is IRS form 941s that provide annual employment and payroll at the EIN-level which can cover multiple establishments. The Census uses an imputation model to allocate these EIN-level measures across establishments. The Census then receives establishment-level information from the Company Organization Survey (COS), Economic Censuses, and Annual Survey of Manufacturers (ASM) that provide more accurate measures of establishment-level employment and the timing of establishment exit but are not conducted for all establishments annually (e.g., the Economic Census is only conducted every five years). See *Chow et al. (2021)* for more details. Consequently, we view our five- and ten-year outcomes as nice checks that these allocation issues are not biasing our overall results or the cause of some of the heterogeneous effects that we observe.

- **Employment**: total number of employees who received wages or other compensation during the pay period that included March 12th.

- **Payroll**: total “wages, tips, and other compensation” for employees over the entire year.
• **Establishment Survival**: indicator for whether the establishment has positive employment for at least one year in the future and in the past. Consequently, an establishment that has 50 employees one year, 0 employees the next, and 50 employees the following year would be defined as a “survivor” in the intermittent year. Since the LBD only measures March 12 employment, these establishments could be true survivors (e.g., seasonal businesses).

**Plant-level TFP from the Annual Survey of Manufacturers** We define plant-level productivity using inputs and outputs from the Annual Survey of Manufacturers (ASM) and TFP measures calculated by Foster et al. (2016). To classify each election into different terciles of the plant productivity distribution, we first take all ASM observations, with and without union elections, with non-missing TFP and calculate year by NAICS 6 industry TFP percentiles. For each of our manufacturing union elections, we then assign the election the plant’s most recent TFP percentile in the previous five years (e.g., if the establishment was sampled by the ASM in year $E_i - 2$ but not $E_i - 1$, we assign the establishment its $E_i - 2$ productivity rank). Based on the election observations with defined TFP, we then classify the elections into within-year terciles based on these rankings.

**Matching Elections, Contracts, and LBD observations** Our data on unionization elections and contract notices contain information on the name and location of the employer, but no unique identifiers (like EIN) that would directly link the establishments to administrative Census data firms. We instead use a fuzzy-matching algorithm to link each establishment to its corresponding Census record from the Standard Statistical Establishment List/Business Register based on the name and geographic similarity of establishments. Our algorithm is based upon the soft-TFIDF approach used by Kline et al. (2019) to fuzz-match firm names from patent records to tax files from the U.S. Treasury; we additionally augment their approach to exploit the additional data and institutional features of our setting.

The standard soft-TFIDF algorithm computes a match score between two firm names that is increasing in their similarity. The algorithm is particularly suitable for our application since it overweights any similarities in uncommon words between the two names and discounts similarities in common words. We evaluate word similarity using the Jaro-Winkler distance, which accommodates misspellings and other variations. For more details on the soft-TFIDF algorithm, see Kline et al. (2019). We first standardize establishment names, and then we compute soft-TFIDF string distances for all election/contract-Census pairs that share at least one word in common.

Although it’s possible to match the unionization records to the Census data based on employer name similarity alone, the procedure is likely to generate false establishment matches (especially given that establishments at multiunit firms may all share the same name, like “CVS” or “Starbucks”). As a result, we incorporate geography information in a flexible manner by geocoding the location of each establishment. When we do not have an address available or we cannot geocode the address, we geocode a higher level of aggregation (e.g., zip code or city). Our final match score between an establishment in the unionization records and an establishment in the Census is then the weighted
mean of their name similarity (based on the soft-TFIDF algorithm) and their geographic proximity (based on the geocodes).

The matching algorithm has several tuning parameters that determine the relative weights placed on each component of the final match score. We use details about our institutional setting to optimize these parameters in a principled manner. We first optimize the soft-TFIDF parameters by matching each election record to at most one contract record. We then choose the string distance parameters that maximize the discontinuity in the likelihood that an election record has a matching contract record across the 50% vote share threshold. We then directly calculate the probability that an election record was matched correctly to a Census record (as a function of the records’ match score) by comparing the number of recorded votes in the election to the number of workers at the Census establishment. For a matched set of records with match score $s$, we define the average likelihood that the matched Census employment is at least as high as the number of recorded votes $m(s)$. On the other hand, the likelihood that the employment at random Census establishment is at least as high as the number of recorded votes is $m$. We assume records where the name and geographic location match exactly are "true" matches, which correspondingly allows us to estimate that a pair of records with a match score of $s$ is matched correctly with probability:

$$p(s) = \frac{m(s) - m}{m(1) - m}. \tag{A1}$$

We include all record matches where the correct match probability $p(s)$ is at least 75%, and we select the geography weight that maximizes the number of elections that are matched in this process. We then use the same parameters to also match contract notices to the Census records.

Although our analysis only considers elections from 1981-2005, we use the earlier election data to restrict the sample to first-time elections at each establishment.
Appendix References

