Unionization, Employer Opposition, and Establishment Closure*

Sean Wang[†] Samuel Young[‡]

January 31st, 2024

Abstract

We study the effect of private-sector unionization on establishment employment and survival. Our empirical strategy extends standard difference-in-differences techniques with regression discontinuity extrapolation methods. 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 their overall opposition to unions. We support these new explanations by showing that firms shift production away from newly unionized establishments and that the negative effects are largest when the firm is likely more opposed to the union.

JEL Codes: J23, J31, J38, J50, J53, J58, J83

*We are grateful to Daron Acemoglu, David Autor, and Simon Jäger for guidance and advice throughout this project. We thank Josh Angrist, Jon Cohen, Emin Dinlersoz, David Hughes, Henry Hyatt, Andrew Joung, Sylvia Klosin, Tom Kochan, Felix Koenig, Mike Piore, Frank Schilbach, Garima Sharma, Aaron Sojourner, Martha Stinson, Martina Uccioli, John Van Reenen, Kirk White, Michael Wong, and Josef Zweimüller and participants at several seminars for helpful comments. We thank Stephanie Bailey, Jim Davis, and Nathan Ramsey for their 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) and by the Washington Center for Equitable Growth. 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, CBDRB-FY22-P2389-R9358, and CBDRB-FY23-CE-005).*
Emails: sean.y.wang@census.gov and sgyoung@asu.edu.

[†]Center for Economic Studies, U.S. Census Bureau

[‡]Arizona State University, W.P. Carey School of Business

1 Introduction

Union elections in the U.S. are extremely contentious. Employers frequently threaten to close establishments if they unionize, and surveys suggest that some follow through on these threats (Bronfenbrenner, 1996). The standard economic explanation for why establishments may close after unionization is that unions make them unprofitable by increasing labor costs or implementing other workplace changes. However, this explanation raises several puzzles. For instance, existing economics research has found little evidence that successful union elections substantially increase wages or decrease productivity.[1]

Additionally, two commonly discussed reasons why unionization may lead to establishment closure are absent from this standard framework. First, some firms can avoid working with new unions rather than needing to shut down entirely (e.g., by reallocating production from unionized to non-unionized establishments). For example, from 2011–2021, Boeing moved the production of all 787 airplanes from a unionized plant in Washington state to a non-union plant in South Carolina to avoid the union representing the Washington workers (Greenhouse, 2011). Second, the degree to which a firm opposes the union could directly affect the consequences of unionization.[2] For example, during a 2017 campaign to unionize the news website Gothamist, the owner stated, "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 unionized, the owner shut down the business (Wamsley, 2017). These examples show that firms' ability to avoid unions and their overall opposition to unions could also explain why unionization may decrease employment and survival.

In this paper, we make two contributions. First, we use a novel research design to analyze how unionization affects establishment employment and survival. We show that successful union elections, including ones away from the 50% winning threshold, lead to large long-run decreases in employment and survival. Second, we test whether a firm's ability to avoid unions or their general opposition to unions helps explain these effects. We find that unionization has the most negative effects when the firm can shift production away from the newly unionized establishment and when it is initially most opposed to unionization. These findings suggest that the standard *establishment-level* framework neglects the role of firms' predisposition and ability to avoid unions as an explanation for unionization's negative effects.

Our setting is around 27,000 U.S. private-sector union elections from 1981–2005 through the National Labor Relations Board. We link these elections to administrative Census data on establishment employment, survival, and productivity. Our empirical strategy adapts regression discontinuity (RD) extrapolation methods to a difference-in-differences framework. This strategy allows us to estimate the effects of successful union elections, including larger margin-of-victory elections. It also avoids biases from only comparing very close elections, which are subject to

[1] See, for example, Frandsen (2021); DiNardo and Lee (2004); Freeman and Kleiner (1990b) on wages and Dube et al. (2016); Sojourner et al. (2015) on productivity. See Appendix C for a broader summary of this literature.

[2] For instance, establishment closures could be driven by owners' or managers' dislike of working with unions rather than unions' direct economic costs (Foulkes, 1980; Leonard, 1992). Alternatively, greater opposition to the union could generate adversarial labor relations that lead to a decline in work conditions (Krueger and Mas, 2004).

vote-share manipulation.

We find that unionization decreases establishment employment, primarily by lowering the likelihood of survival. We estimate a five-year survival effect of four percentage points (pct. pts.) relative to an 82% baseline survival rate. We also find bigger employment declines from larger margin-of-victory elections. Finally, we document significant effect heterogeneity across industries. In the service sector, we find small and sometimes insignificant effects. Instead, the overall negative effects are driven by elections in manufacturing and other blue-collar and industrial sectors.[3] For example, the ten-year survival effect for manufacturing elections is eight pct. pts.

Next, we test whether firms' ability to avoid working with unions or their general opposition to unions helps explain the negative effects we document. For this analysis, we focus on manufacturing elections for three reasons. First, we have better data to test specific hypotheses. Second, it is the largest sector with substantial negative effects. Third, many manufacturing firms produce tradable products, which makes it easier to shift production across establishments.

Our first new hypothesis for these effects is that some firms can avoid working with new unions by shifting production from a newly unionized establishment to their other establishments.[4] To test this, we first estimate whether the effects of unionization are larger at establishments that are part of multi-establishment (MU) firms than at single-establishment (SU) firms. We find significantly larger employment and survival decreases at MU firms. For example, the ten-year survival effects are 12 pct. pts. versus 3 pct. pts. at MU and SU firms, respectively. This heterogeneity is consistent with MU firms avoiding new unions by shifting production to their other establishments.

Next, we more directly test for production shifting after successful elections. Specifically, following successful versus unsuccessful elections at MU firms, we compare the employment growth of the firms' *other* establishments. When we focus on establishments in the same three-digit NAICS industry as the election establishment, we find significantly higher employment growth for the other establishments at firms with successful elections. These same-industry establishments produce similar products to the election establishment, which makes production shifting easier. However, these effects are insignificant five years after an election. Both pieces of evidence support firms avoiding unions through production shifting as one explanation for unionization's negative effect on employment.

Our second hypothesis is that the negative effects of unionization are greater when firms are more opposed to the union. To test this, we estimate treatment effect heterogeneity using two proxies for firms' opposition. First, we estimate effects separately for MU firms with and without any *other* unionized establishments. Survey evidence indicates that firms with lower unionization rates would more "vigorously resist dealing with unions" (Freedman, 1979). Additionally, similar to Selten (1978)'s "chain store paradox," a non-unionized firm might close a newly unionized establishment to convey an aggressive stance on unions, even if it would not be economically rational to close the

[3]Examples of service-sector elections include hospitals, nursing homes, grocery stores, and janitors. The "other" industry group includes transportation, warehouse, and construction elections.

[4]Establishments are distinct locations where employees work. Firms are groups of establishments under the same ownership. Union elections in the U.S. generally occur at the establishment level.

establishment when considered in isolation.[5] Supporting this hypothesis, we find significantly larger long-run employment and survival declines from successful elections at non-unionized firms than unionized firms.

Our second proxy for firms' opposition to the union is the amount of delay time during the election process. Election delay time is a proxy for firms' opposition because it is a key way that they attempt to influence elections. For example, Levitt and Conrow (1993) write that the National Labor Relations Act "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 delay time as the number of days between when the union files for the election and the election date. We estimate separate treatment effects for elections with shorter versus longer delays and find significantly larger employment and survival decreases from longer delay elections. For example, the ten-year survival effect for the top tercile of election delay time is 14 pct. pts. versus 5 pct. pts. for the bottom tercile.

Finally, we test for effect heterogeneity by baseline establishment productivity. If the survival decreases are driven by unions increasing wages or lowering productivity, many theories of firm dynamics predict larger survival declines for lower-productivity establishments. However, we do not find significant differences across establishments with different baseline total-factor productivity, calculated from the Annual Survey of Manufactures. This lack of productivity heterogeneity is more consistent with our new hypotheses than the conventional wage or productivity explanations.

Overall, this heterogeneity analysis shows that unionization has the most negative effects when employers can avoid the union or are initially the most opposed to the union. This evidence can help reconcile the large negative effects on employment and survival we document with previous research on unionization, which has had difficulty finding significant wage increases or productivity declines from the same elections (see Appendix C). First, our production-shifting evidence helps resolve the puzzle because even small wage or productivity effects could lead to large survival declines if firms can cheaply shift production across establishments. Second, one interpretation of our employer opposition evidence is that the overall negative effects are driven by managers' or owners' dislike of working with unions rather than unions' direct economic costs. This interpretation is also consistent with the lack of effect heterogeneity by baseline establishment productivity.[6]

However, we cannot rule out that our proxies for employer opposition reflect rational expectations of unions' economic costs or the hostility of labor relations between the union and firm. This rational future cost explanation is consistent with research by Lee and Mas (2012) and Knepper (2020). While we do not measure these direct costs in our setting, our avoidance and opposition evidence suggest that the effects of unionization on employment and survival may substantially overstate these costs. A fruitful area of future research is unpacking why some firms vigorously oppose unions and the accuracy of their beliefs about the direct costs of unions.

We next summarize our econometric methodology. Our empirical strategy combines features of

[5]One reason to respond excessively to the first election is to prevent unionization from spreading across the firm.

[6]Additionally, Freedman (1979) and Bronfenbrenner (2001) provide survey evidence showing that the firms most opposed to unions were not the firms where unions would have likely imposed the largest costs.

RD and panel data methods that have previously been used to analyze union elections.[7] Although RD designs generally have strong internal validity, they have disadvantages in this setting. First, there is substantial manipulation around the 50% vote-share threshold (Frandsen, 2017). Second, the effects of close elections may be different from elections with larger margins of support. To address these disadvantages, we implement a difference-in-differences (DiD) design that compares establishments that had successful versus unsuccessful union elections. However, we only include elections within a 30 pct. pt. bandwidth around the 50% threshold. This allows us to avoid only comparing very close elections while also excluding less comparable elections with extreme vote shares. 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 show that establishments with winning versus losing elections had similar conditional pre-election employment and payroll growth rates for up to ten years before the elections.

Next, we test additional implications of our identifying assumption that are only possible since we observe election vote shares. These checks extend tests from the RD extrapolation literature to panel-data settings (Angrist and Rokkanen, 2015; Bennett, 2020).[8] They can be implemented in other DiD analyses where the "forcing variable" is observed. First, we show that the similarity in *pre-election* employment growth rates holds between finer vote-share groups. For example, pre-election employment growth rates were statistically indistinguishable between elections with 50–60% and 60–70% vote shares. Second, we show that establishments' *post-election* employment growth and survival were similar between *losing elections* with different vote shares. If our treatment effects were biased by contemporaneous shocks correlated with vote shares, we would also expect these shocks to cause differences between the outcomes at establishments with losing elections with different vote shares. These results support our identifying assumption by showing that it holds for several subsets of observations where we observe untreated potential outcomes.

Finally, we confirm our main results using two alternative identification strategies. First, we leverage multiple elections occurring at different establishments within the same firm. Specifically, we compare winning versus losing elections that occurred *at the same firm* and in the same year, ensuring that these establishments experienced the same firm-level shocks. For our second alternative strategy, we compare elections within a narrower 40–60% vote-share bandwidth. Both strategies yield no employment growth pre-trends and large negative treatment effects, even without conditioning on any baseline covariates.

Our overall employment and survival estimates contribute to the literature on the effects of unionization in the U.S. Due to our different empirical strategies, our estimates complement Frandsen (2021)'s RD estimates of short-run employment decreases and his suggestive evidence of

[7]See DiNardo and Lee (2004); Sojourner et al. (2015); Knepper (2020) for RD analyses and Freeman and Kleiner (1990b); LaLonde et al. (1996); Lee and Mas (2012); Dube et al. (2016) for panel data analyses. Frandsen (2021) also combines these methods by implementing an RD design on first-differenced outcomes.

[8]Lee and Mas (2012), Frandsen (2021), and Sojourner and Yang (2022) present pre- and post-election outcomes *across the vote-share distribution* but do not use these estimates as formal tests of their identifying assumptions.

negative survival effects.[9] Our findings are also consistent with LaLonde et al. (1996) and Sojourner et al. (2015), who find that unionization decreases employment in specific sectors.[10] However, our results contrast with DiNardo and Lee (2004)'s null effects for survival and employment and other research that finds no survival effects (Freeman and Kleiner, 1999). These differences may be due to our use of higher-quality establishment survival data. Finally, although long hypothesized, we provide the first evidence that the effects of unionization on survival vary across sectors.

Our evidence supporting the union avoidance and employer opposition hypotheses is novel relative to prior economics research but consistent with previous labor relations research. For example, Bronfenbrenner (2000, 2001) report similar results from a survey of union organizers in the 1990s. She finds survival declines of 12 pct. pts. following successful elections. She also finds that establishment-closing threats were more common in the types of elections where we find larger survival effects (e.g., manufacturing and MU firms). Additionally, our evidence that employers who were more opposed to unions were more likely to close unionized establishments adds to the literature on anti-union firms' union avoidance tactics (Freeman and Kleiner, 1990a; Kleiner, 2001). Finally, our production-shifting evidence is consistent with firms becoming less unionized by shifting investment to non-union establishments (Verma, 1985; Kochan et al., 1986a) and with Giroud and Mueller (2017) and Guo (2023)'s finding that MU firms shift production away from high-tax locations.

Our paper is structured as follows. We describe union elections in Section 2 and our data in Section 3. Section 4 discusses our empirical strategy. Section 5 presents estimates of the effects of unionization on employment and survival. Section 6 provides evidence supporting our avoidance and opposition hypotheses. Section 7 discusses our results.

2 Unionization through NLRB Elections

The National Labor Relations Act (NLRA) guarantees most U.S. private-sector workers the right to collective bargaining. Under the NLRA, an employer is required to bargain over the conditions of employment with any union that represents a group of its workers.[11] Bargaining generally occurs at the establishment level (Traxler, 1994). During negotiations, the union may go on strike to pressure the employer. The NLRA also created the National Labor Relations Board (NLRB), a quasi-judicial agency that administers union elections and adjudicates unfair labor practice charges. The current U.S. policy debate about organized labor focuses on increasing representation at non-union establishments (e.g., the *PRO Act*). Our results speak directly to the potential consequences of these efforts.

The primary way for private-sector workers to gain union representation is a secret-ballot

[9] His survival estimates are differences in survival probabilities around the 50% threshold, and he states, "a causal interpretation of the differences in survival probability should be made with caution" due to vote-share manipulation.

[10] LaLonde et al. (1996) analyzed manufacturing elections using a DiD design. However, they do not analyze the effect on survival, which makes interpreting the results conditional on survival difficult. Additionally, due to a smaller sample, their pre-trend estimates are imprecise, making it hard to evaluate the parallel trends assumption.

[11] These include wage and non-wage compensation and promotion, grievance, and layoff policies.

NLRB election.[12] An 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 signed indicating union support by workers in the "bargaining unit" (i.e., the workers the union would represent). A bargaining unit usually only contains workers at a single establishment but can range from only workers in one occupation (e.g., delivery truck drivers) to all non-managerial employees. After gathering signatures from at least 30% of the bargaining unit, the union files an election petition. The NLRB then validates the signatures, resolves bargaining unit disputes, and schedules an election. After the petition is filed, employers frequently try to delay the election to reduce union support (e.g., contest the bargaining unit composition) (Levitt and Conrow, 1993; McAlevey, 2020).

Unions and employers often campaign before the election. Union organizers and pro-union workers campaign by speaking with workers at work or during house calls, publicly showing solidarity (e.g., rallies), and enlisting community support (Bronfenbrenner and Juravich, 1998). Employers use tactics like holding one-on-one meetings with supervisors, requiring employees to attend "captive audience meetings," and hiring "union avoidance" consultants (Logan, 2002). Finally, although there are legal restrictions on firing pro-union workers and threatening to close establishments, there is evidence that firms use these tactics (Weiler, 1983; Schmitt and Zipperer, 2009).

If a majority of workers votes for the union, the union is certified to represent the bargaining unit. Afterward, the employer must bargain "in good faith" with the union, but they are not required to reach an agreement.[13] If no contract is reached within a year of certification, the employees can vote out the union in a *decertification* election.

The NLRA also limits whether firms can close newly unionized establishments. Generally, establishment closures violate the NLRA if they are motivated by "anti-union sentiment." Instead, closures motivated by "economic reasons" do not violate the NLRA. In actual cases, the NLRB considers whether the firm's statements suggest an anti-union sentiment, whether the firm planned the closure before the election, and the closure's timing relative to the election (Munger et al., 1988). Finally, closing an *entire firm* rather than a single establishment typically does not violate the NLRA, regardless of the motivation.

Selection into Union Elections Since our empirical strategy compares winning and losing elections, we next review the literature on selection into holding and winning elections. This literature motivates which baseline characteristics we condition on and our additional tests of the identifying assumption. 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 held elections.

[12] While NLRB elections are the primary method for private-sector workers to unionize, unionization can occur without an NLRB election. First, the NLRA does not cover all workers (GAO, 2002). Some workers lack collective bargaining rights (e.g., some small business, domestic, and agricultural workers). Others have collective bargaining rights but are not covered by the NLRA (e.g., airline workers or public-sector workers). Second, covered workers can also gain representation through voluntary "card check" recognition. However, card check is much less common than elections. Schmitt and Zipperer (2009) estimate that from 1998–2003, 60% of new union recognition occurred through NLRB elections, but assume that before then, 90% of organizing occurred through elections.

[13] In a review, CRS (2013) found that 56–85% of successful elections result in first contracts during our timeframe.

Workers, employers, and other factors could all influence election outcomes. For our empirical strategy, the concern is that election vote shares may be related to future establishment productivity changes. For example, workers who expect their establishment to become more productive and have more rents to share may be more likely to vote for a union. Alternatively, firms that expect to become more productive may campaign harder against unions, which would generate a negative bias.

Prior research on the determinants of union election outcomes finds that all these factors play a role. The most consistent finding is higher win rates for smaller bargaining units (Heneman and Sandver, 1983; Farber, 2001). Win rates also vary substantially across industries (Bronfenbrenner, 2002). This motivates our first set of controls that just include establishments' baseline employment and industry. Regarding the influence of employer versus union campaigns, Bronfenbrenner (1997) finds that "union tactic variables explain more of the variance in election outcomes than any other group," including employer tactics or characteristics. Yet, other research finds that strong anti-union campaigns are associated with lower win rates (Freeman and Medoff, 1984). To assess whether employers' campaigns lead to a negative bias between vote shares and establishment growth, we implement multiple tests of how vote shares are related to pre- and post-election outcomes.

Motivation for Getting Away from the RD Threshold An advantage of our empirical strategy is that it does not only compare elections around the 50% threshold. One motivation for this is the non-random sorting of elections around this threshold (i.e., "vote-share manipulation"). Figure 1 Panel A plots the vote-share distribution of elections with at least 50 votes and shows a missing mass of barely winning elections (Frandsen, 2017).[14] Frandsen (2021) shows that this also leads to large differences in pre-election establishment characteristics across the threshold (e.g., 13–22% employment level differences).

Another motivation for not just comparing close elections is that the treatment effect of unionization may depend on the election vote share. For example, Lee and Mas (2012) only found negative stock price effects of unionization for higher margin-of-victory elections. One potential reason for this heterogeneity is that close elections are often followed by delays before bargaining starts (e.g., debates about challenged votes). Figure 1 Panel B shows this by plotting the average number of days between the election date and the case-closing date (i.e., when the union is certified). The figure shows a large increase in bargaining delay for very close elections (e.g., the median (mean) for 51% vote-share elections is 118 (223) days versus 11 (57) for 60% vote-share elections). Since delays can dampen unions' bargaining power, Panel B suggests that the effects of union elections may differ between very close elections and those with even slightly higher vote shares.

Additionally, unions that are certified in very close elections are more likely to face a *decertification* election soon afterward. Figure 1 Panel C plots the probability that a certification election will be followed by a decertification election at the same establishment in the next five years. It

[14]Figure 1 only includes elections with at least 50 votes because the discreteness of the running variable biases conventional manipulation tests for low-vote elections. Frandsen (2017) provides evidence of manipulation using tests for discrete running variables. Figure A1 shows the vote-share histogram that includes all elections in our sample.

shows that more than 12% of very close winning certification elections experience a decertification election compared to less than five percent of higher vote-share elections. Overall, Panels B and C show that proxies for unions' bargaining power increase in the election vote share, suggesting that the effects of unionization also differ along this margin.

3 Election, Contract, and Establishment Data

We combine union election and contract data with administrative data from the U.S. Census Bureau. These data are ideal for studying union elections (DiNardo and Lee, 2004; Dinlersoz and Greenwood, 2016; Frandsen, 2021). First, they contain the universe of *establishments*, the level at which most elections occur. Second, the Census constructs high-quality longitudinal links that allow us to separate real establishment exits from spurious exits due to administrative or ownership changes (Haltiwanger et al., 2013). See Appendix D for more details on the data, sample selection, and matching.

NLRB Election Data We combine data from multiple sources to construct a comprehensive dataset of union elections from 1962 to 2018.[15] The data contain election vote counts that we use to define treatment and employers' names and addresses that we use for matching to the Census data. Additionally, they include the election filing date, election date, and case-closing date. We define the treatment year based on the election filing date because this is the earliest observed date for each election. We also use these dates to define "election delay time" (i.e., the number of days between filing and holding an election).

FMCS Contract Data To measure whether an establishment is covered by any collective bargaining agreement (CBA), we use contract notice data from the Federal Mediation and Conciliation Service (FMCS) from 1984–2019. The data includes notices of initial contract negotiations and renegotiations for existing contracts. These "notices of bargaining" are provided to the FMCS so that it can be ready to provide mediation. Although filing is legally incentivized, underreporting is likely. We use these data to measure whether an election establishment has other workers already covered by a different CBA and whether the establishment's firm has other unionized establishments.

Employment, Payroll, and Survival Data Our primary outcomes are from 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; Chow et al., 2021). The LBD employment measure is the total number of employees on March 12th of each year. The payroll measure is employees' total "wages, tips, and other compensation" over the entire year. Consequently, in an election year, we expect larger effects on payroll than employment. The data also contain high-quality longitudinal establishment IDs that identify the same establishments over time, even across ownership changes. We use these IDs to define establishment survival based on

[15] Appendix Figure A2 shows that we have a similar number of cases each year as in the NLRB's annual reports.

the last year of non-zero employment. Finally, we use Fort and Klimek (2016)'s 2012 NAICS codes to classify each establishment into time-consistent industry codes.

We address potential biases from how the Census calculates employment at *multi-establishment (MU) firms* by focusing on long-run outcomes. Although the LBD is at the establishment level, employment and payroll data are sometimes received at the more aggregated EIN level. These aggregate measures are allocated proportionately across establishments based on their past values. Consequently, if a unionized establishment at an MU firm decreases its employment, some of this decrease may be allocated to the firm's other establishments, underestimating the effect of unionization. To address this bias, we focus on five- and ten-year outcomes since the Census receives *establishment-level* payroll and employment data at least every five years.

Sample Selection and Matching Before matching elections to Census data, we restrict the sample to focus on elections that likely shifted establishments' union status. Appendix Table A1 shows how these restrictions affect the sample size. First, we restrict to elections from 1981–2005. We end the sample in 2005 to create a balanced five-year pre-period and ten-year post-period for each observation (the version of the LBD we use is available from 1976–2015). Second, we drop non-representation cases (e.g., decertification elections). Third, we drop elections involving multiple unions. These are often "union raids" involving incumbent unions and may only change *which* union represents the workers (Sandver and Ready, 1998). Fourth, we drop elections with fewer than six eligible voters to ensure a non-trivial increase in union representation.

Next, we match each election to an establishment in the LBD. First, for each election and LBD observation, we calculate a weighted average of the *Soft TF-IDF* distance between the employer names and the geographic distance between geocoded addresses. We match each election to the LBD establishment with the highest match score above a minimum threshold. This procedure yields a 70% match rate. We use the same procedure for each FMCS notice.

We further restrict the matched sample based on the requirements of our empirical strategy. First, we only keep the first election at each establishment. Next, we drop elections at establishments less than three years old. Since a key test of our identifying assumption is that the outcomes for winning and losing elections evolved similarly beforehand, we exclude observations where we cannot evaluate this for at least three years. To keep our sample the same across specifications, we require non-zero payroll and employment one year before the election. This results in a sample of around 27,000 elections (see Appendix Table A1).

Finally, for much of our analysis, we restrict to 20–80% vote-share elections. This decreases our sample to 19,000 elections. A motivation for this restriction is that the extreme vote-share elections have some different ex-ante characteristics than other elections.[16] Consequently, the establishments with these elections appear to have systematically different pre-election trends than establishments

[16] Figure 2 in Frandsen (2021) shows relatively similar pre-election employment and payroll for 20–80% elections, but divergences for the other elections. Similarly, Appendix Figure A3 shows that the 0–20% elections are less likely to be non-first-elections at the establishment and that 80–100% elections are more likely to be multi-union elections. Although we exclude non-first and multi-union elections from our analysis, this evidence suggests that these elections, where unions have nearly unanimous support or opposition, may also differ from the other elections on other unobservable dimensions.

with other elections. As a result, our identifying assumptions are less likely to hold for these elections. However, our results do not rely on the particular bandwidth choices of 20% and 80%. For example, Figures 5 and 6 show that the overall estimates are not driven by the 20–30% or 70–80% elections. Similarly, Tables A3, A9, and A11 show that our heterogeneity estimates are qualitatively the same with a 30–70% bandwidth.

Table 1 presents summary statistics for our sample. It confirms the characteristics of winning elections described in Section 2. Winning election establishments are smaller, less likely to be part of MU firms, and more likely to have another unionized bargaining unit. However, the differences are small for average payroll and establishment age.

4 Empirical Strategy and Identifying Assumptions

Our research design combines difference-in-differences (DiD) techniques with tests of our identifying assumption from the RD extrapolation literature. Specifically, since we observe vote shares that determine treatment assignment, we can assess several testable implications of this assumption that are not possible in standard DiD settings.

Potential Outcomes To fix ideas, consider all establishments, i, that held an election in 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%[17]

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

An establishment's non-unionized potential outcome is Y_{it}^0. Its unionized potential outcome is $Y_{it}^E(V)$, which may depend on its cohort E and vote share V. This allows for dynamic treatment effects and treatment effect heterogeneity 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

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

Our estimand of interest is the treatment effect n years after a successful election with vote share V

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

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

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

[17]This definition assumes that treatment is absorbing (i.e., $D_{it} = 1 \Rightarrow D_{it'} = 1 \ \forall \ t' > t$). 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 always correspond with union representation or a contract (e.g., future decertification elections).

where γ_i are establishment fixed effects (FEs) and α_t are year FEs.[18] The coefficients of interest, δ_n, capture the average, dynamic treatment effects of a successful union election. X_i are baseline, one year before the election, establishment characteristics whose coefficients vary with event time n.

Identifying Assumption Our identifying assumption is conditional parallel trends *by vote share*

$$\mathrm{E}\left[Y_{it}^0 - Y_{it-1}^0 | X_i, V_i\right] = \mathrm{E}\left[Y_{it}^0 - Y_{it-1}^0 | X_i\right]. \tag{5}$$

Intuitively, we assume that outcomes at establishments with different election vote shares but the same baseline characteristics would have followed parallel trends had no election occurred. There are several things to note about this assumption. First, it does not restrict selection *into* elections (e.g., organizers targeting productive establishments) or *selection on gains* based on the effects of unionization (e.g., workers 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 treated and control observations. Yet, this stronger assumption yields richer testable implications. Additionally, we do not need to impose this assumption across the entire vote-share distribution. Instead, we can only assume it holds for some bandwidth around the 50% threshold and assess its testable implications within this bandwidth. This illustrates how our strategy allows us to get away from only comparing elections right around the threshold without assuming parallel trends between *all* elections. Finally, vote shares may be influenced by workers, employers, and other factors that could lead to violations of this assumption. This motivates our conditioning on specific baseline covariates and assessing testable implications of the assumption to provide reassurance that such selection does not bias our results.

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

Testable Implications Our identifying assumption yields several testable implications. Intuitively, we observe untreated potential outcomes, Y_{it}^0, for many observations and can test whether equation 5 holds for different subsets of these observations.

The first testable implication of equation 5 is that there should be conditional parallel trends in **pre-election** outcomes across **all vote shares** included in the sample

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

This test nests the standard DiD pre-trends test between winning versus losing elections on average. Moreover, we can test for similar pre-trends between finer vote-share groups. For example, we

[18] We exclude establishment FEs for outcomes that are identical for all establishments one year before the election (e.g., survival and DHS growth rates). For DHS growth rates, we capture the time-invariant component by differencing relative to $t = E_i - 1$. We include establishment FEs for log outcomes.

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 RD identification away from the threshold. They argue that conditional mean independence of untreated potential outcomes and the running variable for a given bandwidth around the RD threshold is strong support for parallel trends holding within that bandwidth.

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

$$\mathrm{E}\left[Y_{it} - Y_{it-1} | X_i, V_i\right] = \mathrm{E}\left[Y_{it} - Y_{it-1} | X_i\right] \text{ for all } t \geq E_i \text{ \& } V_i \leq .5. \tag{7}$$

To implement this test, we estimate whether post-election outcomes are different between losing elections with different vote shares (e.g., compare conditional post-election survival rates for 30–40% versus 40–50% elections). This test addresses the concern that 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 establishments with losing elections with different vote shares.

Figure 2 illustrates these testable implications. It plots average hypothetical outcomes two years before an election, $Y_{i,-2}$ and $Y_{i,-1}$, and one year after, $Y_{i,1}$, by vote share. Testing parallel *pre-trends* (e.g., equation 6) by vote share corresponds to comparing the distance between the $Y_{i,-2}$ and $Y_{i,-1}$ lines. Testing parallel *post-trends for losing elections* (e.g., equation 7) corresponds to comparing $Y_{i,-1}$ and $Y_{i,1}$ for losing elections.

Pooling all Cohorts Our sample includes all election cohorts from 1981–2005. To estimate the effect across all cohorts, we pool these elections and estimate

$$Y_{it} = \gamma_i + \alpha_{t,E_i} + \sum_n \delta_n \cdot \mathbb{1}[t - E_i = n] \times \mathbb{1}[V_i > .5] + X_i' \beta_{n,E_i} + \varepsilon_{it}. \tag{8}$$

This specification is the same as the single-cohort specification (equation 4), except that the year FE and baseline control coefficients can now vary by cohort (i.e., α_{t,E_i} and β_{n,E_i} have E_i subscripts). The motivation for this flexibility is that by interacting these variables by cohort, our specification is analogous to the "stacked regression" approach to DiD settings with treatment time variation (Cengiz et al., 2019). Consequently, we avoid the negative weight issues that arise from heterogeneous, cohort-specific treatment effects in this setting (Sun and Abraham, 2021; Goodman-Bacon, 2021; de Chaisemartin and D'Haultfoeuille, 2020).[19] Intuitively, our estimates come from only comparing winning and losing elections *within the same cohort* rather than making any "forbidden comparisons" between winning elections in different cohorts. An additional benefit of this specification is that we only need to assume that our identifying assumption in equation 5 holds *within each cohort*.[20]

[19]Additionally, we test for negative weights on each cohort treatment effect in the specification we implement with controls using Sun and Abraham (2021)'s `eventstudyweights` package.

[20]With multiple cohorts, our identifying assumption is $\mathrm{E}\left[Y^0_{it} - Y^0_{it-1} | X_i, E_i, V_i\right] = \mathrm{E}\left[Y^0_{it} - Y^0_{it-1} | X_i, E_i\right]$.

However, we show that our results are qualitatively the same with and without interacting the baseline controls by cohort. Finally, we cluster standard errors at the *firm* level.[21]

Establishment-Level Controls To account for observable determinants of election outcomes, we control for progressively richer establishment 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 only employment and three-digit NAICS industry-by-year controls.[22] The motivation for starting with these covariates is that they are among the strongest predictors of election victory (see Section 2), and they are key determinants of establishment growth and survival dynamics (Dunne et al., 1989; Haltiwanger et al., 2013). Our second, *flexible controls* specification includes other establishment characteristics and allows the control coefficients to vary by cohort (i.e., we interact the baseline controls with year of election). Specifically, we add other characteristics in the LBD (payroll, establishment age, and single/multi-establishment firm status) and an indicator for whether we observe a previous FMCS contract at the establishment (i.e., another bargaining unit already unionized at the establishment).[23] The cohort interactions in this control specification result in the analog to the "stacked regression" approach discussed previously. We show, however, that all our results are robust to only including the pooled 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

$$G_{it} = 2 \times \frac{Y_{i,t} - Y_{i,t=E_i-1}}{Y_{i,t} + Y_{i,t=E_i-1}}. \tag{9}$$

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 estimate of G_{it} could represent either an approximately 20% decline in intensive margin employment with no survival effects or a ten pct. pt. decrease in the likelihood of establishment survival. Since the growth rate accommodates exit, we can simultaneously evaluate pre-trends and interpret treatment effects, even if unionization affects establishment survival, which could lead to a selected group of survivors. For this reason, the DHS growth rate is commonly used to analyze firm growth dynamics

[21]This accounts for serial correlation across time and across elections at different *establishments* within a *firm*.

[22]Our 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). For these outcomes, only 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.

[23]The motivation for including the previous contract control is that union elections are more successful when other workers at the same establishment are already unionized (Bronfenbrenner, 2002). 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 the controls in the *flexible controls* specification with our three coarse industry groups (e.g., manufacturing, services, and "other"). This keeps the controls at the same level of granularity for the all industries and manufacturing sample estimates.

(Haltiwanger et al., 2013; Davis et al., 2014; Chodorow-Reich, 2014).

To estimate the effect on survival, we include an indicator for whether the establishment existed at time t as the outcome. We can compare the survival and DHS growth rate effects to assess what share of the DHS effect is *mechanically* due to exit (e.g., $G_{it} = -0.2$ can be explained by a 10 pct. pt. decrease in survival). However, the residual part not explained by exit could be due to either intensive-margin employment changes or selective exit based on employment growth rates.

Finally, we include log employment and payroll as outcomes. The pre-trends for these outcomes are a useful complement to the DHS growth rate pre-trends.[24] However, a challenge with interpreting the post-election effects on log outcomes is that treatment effects on survival can bias comparisons of potentially selected *survivors*. We provide two ways of partially alleviating this concern. First, all specifications with log outcomes include establishment FEs that account for *level* differences between surviving and exiting establishments. Second, the timing of the log outcome versus survival effects suggests intensive margin effects (e.g., large log outcome effects before large survival effects).

Our parallel trends assumption in equation 5 imposes a specific functional form restriction for each outcome (Roth and Sant'Anna, 2023). First, we assume that log employment and payroll would have (conditionally) evolved in parallel, which is theoretically plausible in this setting.[25] Additionally, we test whether the restriction holds in the pre-period (i.e., testing equation 6). For establishment survival, we assume that the survival probabilities between establishments with different election vote shares would have (conditionally) been equal had no elections occurred. We cannot test this assumption in the pre-period since all establishments exist at event-time zero. However, we can test whether the functional form assumption holds between the losing elections with different vote shares (i.e., testing equation 7). The DHS growth rate is approximately a linear combination of the other outcomes, so the previous two functional form assumptions approximately imply parallel trends for DHS.[26]

5 Results: Employment and Survival Effects

We estimate the effects of unionization on establishment employment and survival. We first compare employment growth at establishments with winning and losing elections. Next, we implement the tests of our identifying assumption described in Section 4. Since we later focus on manufacturing, we present our estimates separately for manufacturing elections.

[24] The DHS pre-trends combine intensive and extensive margin employment changes, while the log pre-trends only capture intensive margin changes. However, in specifications where we control for baseline establishment age, the DHS pre-trends will closely approximate pre-trends for log outcomes.

[25] For example, consider two firms with the same Cobb Douglas production function but different baseline TFP 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 payroll and log employment would both evolve in parallel while their levels would diverge.

[26] $\mathrm{E}[\Delta \ln Y^0_{it}|X_i, V_i] = \mathrm{E}[\Delta \ln Y^0_{it}|X_i]$ and $\mathrm{E}[\mathbb{1}[Y^0_{i,t} = 0]|X_i, V_i] = \mathrm{E}[\mathbb{1}[Y^0_{i,t} = 0]|X_i]$ imply $\mathrm{E}[G^0_{it}|X_i, V_i] \approx \mathrm{E}[G^0_{it}|X_i]$.

5.1 Employment and Survival Estimates

We start by estimating establishment employment growth for successful versus unsuccessful elections. Figure 3 plots the δ_n coefficients from the "pooled cohort" specification in equation 8 for 20–80% vote share elections. Panel A plots the estimates for DHS employment growth relative to one year before the election. Panel B includes log employment as the outcome. Both panels include estimates with no controls, only baseline industry and employment controls, and the flexible control specification.

Panels A and B show that establishments with successful elections had similar conditional pre-election growth rates to those with unsuccessful elections, but experienced large employment decreases after the election. The "no control" estimates show that establishments where the union won had relatively slower unconditional pre-election employment growth rates than those where the union lost. However, the "industry + emp ctrls." estimates show that just conditioning on baseline employment and industry yields similar pre-election growth rates for DHS and log employment.[27] Starting one year after the election, this specification shows decreased employment for establishments with successful elections. The effects stabilize about three years after the election. Finally, the "flexible control" specification shows that our pre- and post-election employment growth estimates are qualitatively unaffected when we add these richer controls.

To help interpret the magnitude and timing of these effects, Panel C additionally plots payroll and establishment survival estimates. It includes estimates of DHS employment and payroll growth and establishment survival from the flexible control specification. We find that payroll initially declines faster than employment. This difference could be due to compositional shifts to lower-wage workers or differences in the timing of the payroll versus employment variables (see Section 3). Five years after a successful election, the DHS employment and payroll growth rates are -0.13 and -0.14 lower, respectively (consistent with a 14% decrease in payroll or a seven pct. pt. decrease in survival probability). Appendix Figure A4 presents estimates from the same specification for *log* employment and payroll. These estimates reject five-year, pre-election growth rate differences of more than 3.5% for employment and 3.0% for payroll. In other words, we can rule out that winning establishments grew more than 3.5% faster than losing ones in the five years before the election. Additionally, the figure shows a five-year intensive margin decline in employment of 7%, although we interpret this cautiously given potential biases from selective exit.

The survival estimates in Figure 3 Panel C indicate that most of the decrease in DHS employment and payroll is due to reduced establishment survival. To decompose what share of the DHS effects is from survival, 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 shows how much of the DHS effect is mechanically due to the survival effect (see Section 4). The figure shows a five-year survival effect of four pct. pts. Consequently, about two-thirds of the -0.13 five-year DHS employment estimate is mechanically due to decreased survival. The slower timing of survival

[27] Similarly, Dube et al. (2016) find similar productivity pre-trends for nursing home elections, and Lee and Mas (2012) find similar stock-price pre-trends, which also incorporate expectations of future productivity growth.

versus employment effects may be due to an increased risk of violating the NLRA when immediately closing an establishment after an election.

Given our later focus on manufacturing, Figure 4 presents the same estimates for manufacturing elections. For these elections, we find no detectable employment pre-trends, even with no controls. For example, Panel B shows that we can rule out five-year unconditional employment growth rate differences of more than five percent. One explanation for the lack of pre-trends without controls is that by comparing only manufacturing elections, we account for sector differences captured by controls when we pool all industries. Additionally, for manufacturing, the treatment effects are larger than for all industries (e.g., the five-year DHS employment estimates are -0.17 versus -0.13). We later show that this is because the effects of unionization in the service sector are small.

5.2 Nonparametric Vote-Share Heterogeneity Tests

Next, we provide more evidence that our results are driven by unionization by assessing several testable implications of our identifying assumption. Specifically, we visually show how treatment effects and pre-election trends vary across the vote-share distribution. This also allows us to estimate treatment effect heterogeneity by vote share.

To estimate pre-trends and treatment effects for different parts of the vote-share distribution, we estimate a modified version of our main specification[28]

$$Y_{it} = \alpha_{t,E_i} + \sum_g \sum_n \delta_{g,n} \cdot \mathbb{1}[t - E_i = n] \times \mathbb{1}[V_i \in \mathcal{V}^g] + X_i' \beta_{n,E_i} + \varepsilon_{it}. \tag{10}$$

$\mathcal{V}^g$ are subsets of the vote-share distribution. Specifically, we include eight vote-share 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 all estimates are relative to 20–30% elections. This specification allows us to assess the testable implications of our identifying assumption in Section 4. First, we test whether *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 $V_i \leq .5$ (i.e., testing equation 7). We first present these estimates for manufacturing because our results closely support our identifying assumption, which makes them easier to explain. We then present estimates for all industries where we reject these tests for some outcomes. We find, however, that the violations are driven by elections with exactly 50 % vote shares and discuss potential explanations.

Figure 5 presents estimates from equation 10 for all manufacturing elections. The estimates include our *flexible controls* specification (see the later parametric tests for robustness to alternative controls). Panel A includes DHS employment growth estimates for each vote-share group. For example, the green "five-year pre" estimates show how pre-election employment growth rates differ for each vote-share group relative to elections with 20–30% vote shares. First, the five-, three-, and

[28]Equation 10 omits establishment FEs because we only estimate this specification for outcomes where we do not include these FEs.

two-year pre-trend estimates are similar across almost the entire vote-share distribution (the one exception is 0–20% elections which we exclude from our main analysis). These results support our identifying assumption by showing that the similarity in pre-election employment growth rates holds between much finer vote-share groups. Second, the figure shows that none of the five- and ten-year treatment effect estimates for *losing elections* are significantly different than the estimates for 20–30% elections. This similarity provides reassurance against the concern that future productivity shocks correlated with vote shares bias our main estimates. In that case, we would also expect these shocks to cause outcome differences for losing elections with different vote shares. Finally, the five- and ten-year treatment effect estimates for winning elections are larger for higher vote-share elections (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 for establishment survival. Although we cannot test for survival pre-trends, we can test for parallel trends in survival treatment effects between losing elections with different vote shares. The post-election survival rates for all losing election vote-share groups are not statistically different than for 20–30% elections. For winning elections, however, the figure shows larger long-run survival declines for higher vote-share elections, although the differences are not statistically significant.

Figure 6 presents vote-share heterogeneity estimates when we pool together all industries. Panel A shows that for all 20–80% vote-share groups, we find very similar pre-election DHS employment growth rates. For 0–20% and 80–100% elections, however, we find significantly different pre- and post-election growth rates, which is one reason we exclude these elections from our main analysis. For post-election outcomes, we find similar employment growth rates between 20–30% and 30–40% elections but find somewhat slower employment growth for 40–50% 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 the votes, and there are multiple reasons that these elections are different than elections where the union lost by slightly larger margins.[29] To show this, we estimate the effect for 40–50% elections but exclude elections with exactly 50% vote shares. We find five- and ten-year estimates of -0.015 (SE 0.025) and -0.032 (SE 0.028), respectively. Both estimates are insignificant and economically smaller than the large treatment effects for the neighboring group of 50–60% elections (-0.11 and -0.16). This shows that without the 50% elections, there is no evidence of differential post-election outcomes among losing groups with 20–50% vote shares. Furthermore, Panel B of Figure 6 shows no evidence of differential survival rates between 20–30, 30–40, and 40–50% losing elections, even including the 50% elections.

[29]There are several potential reasons for outcome differences at establishments with 50% vote-share elections. First, due to the discreteness of total votes, elections with 50% vote shares have a small number of total votes cast. Based on the NLRB data, the median (mean) number of voters in 50% vote-share elections is 12 (22) compared to 50 (96) in elections with vote shares in the $[45, 50)$ range. Although our employment controls capture establishment size differences, they do not capture 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).

5.3 Alternative Identification Strategies

Next, we implement two alternative empirical strategies that rely on new identification assumptions. Our first strategy compares winning and losing elections within a narrower vote-share bandwidth. Specifically, we limit the sample to 40–60% elections, excluding 47.5-52.5% elections to account for potential biases from the vote-share manipulation. With this sample, we estimate equation 8 but only include year-by-cohort FEs as controls. Since these winning and losing elections often only differed by a handful of votes, it is plausible that their outcomes would have evolved in parallel, even without conditioning on baseline covariates. Consequently, this strategy combines the logic of RD designs that compare very close elections while also avoiding biases from vote share manipulation.

Our second strategy compares winning and losing elections within the same firm. Specifically, we limit our sample to elections at *firms* with at least one winning and losing election in the same year ($\approx$ 1,000 elections). We then estimate equation 8 but only include firm-by-year-by-cohort FEs as controls so that we identify the impact of unionization by only comparing establishments within the same firm (since DHS is a difference, we account for time-invariant establishment characteristics).[30] Consequently, this strategy accounts for all firm-level differences between winning and losing elections that could bias our results (e.g., different firm-level demand shocks or intensities of anti-union campaigns).

Figure 7 plots DHS employment estimates from both strategies. For reference, we also plot estimates from our main strategy. For both new strategies, we do not detect significant employment pre-trends. This supports our assumption that elections with similar vote shares or elections at the same firm provide a suitable control group for winning elections. Additionally, the strategies yield qualitatively similar negative estimates of the effects of unionization as our baseline strategy. Finally, Appendix Table A2 shows that both strategies also yield significant negative survival effects.[31]

The estimates from these alternative strategies provide further evidence that unionization decreases employment and survival. Additionally, they show this without needing to assume parallel trends *conditional* on baseline controls. Instead, these strategies account for selection into unionization by comparing very close elections or elections within the same firm. However, one notable difference between these estimates and our main results is that these alternative strategies yield wider confidence intervals. This illustrates a trade-off between strategies – while our main strategy relies on a potentially stronger identifying assumption, it gives us more statistical power to estimate effect heterogeneity.

[30]Although this strategy has a potential SUTVA violation, this violation is unlikely to explain the magnitude of the treatment effects we estimate. Specifically, there may be spillover effects from the winning elections to establishments that had the losing elections. However, when we directly estimate these spillover effects in Section 6, the magnitude of the coefficients ranges between 0.03–0.05, much smaller than the treatment effects of up to -0.2 that we find with this alternative identification strategy. Additionally, those spillover estimates are from same-industry, manufacturing establishments. For this strategy, we include non-manufacturing and within-firm establishments in different industries, where we would expect these spillovers to be considerably smaller.

[31]When we implement these alternative strategies for just manufacturing elections, the limited bandwidth estimates are the same sign and significance as our main estimates. Due to a small sample size, however, the within-firm strategy yields imprecise estimates that are not significantly different than zero.

5.4 Industry Heterogeneity Estimates

Next, we separately estimate the effects for different industries and show that the overall effects are driven by elections in manufacturing and other blue-collar and industrial sectors. There are multiple reasons to expect heterogeneity across industries. First, the quality of labor relations may differ across sectors (e.g., strikes were more common in manufacturing, suggesting more adversarial relations). Second, firms in different industries differ in how easily they can "avoid unions." For example, manufacturing firms may avoid working with new unions by shifting production to other establishments. However, this tactic may be difficult in non-tradable industries (e.g., hospitals). Finally, the elasticity of labor demand may differ across industries, which would lead to heterogeneous employment declines in response to identical wage changes.[32]

We first classify elections into three industry groups: manufacturing, services, and other blue-collar and industrial sectors.[33] Over 70% of voters in service-sector elections are in healthcare, security, restaurants, grocery stores, universities, and print media. The "other" category includes agriculture, construction, mining, transportation, warehousing, utilities, and wholesale trade.

To estimate industry heterogeneity, we use the following specification for a categorical heterogeneity variable H_i (e.g., the three industry groups)[34]

$$Y_{it} = \alpha_{t,E_i} + \sum_h \sum_n \delta_{h,n} \cdot \mathbb{1}[t - E_i = n] \times \mathbb{1}[V_i > .5] \times \mathbb{1}[H_i = h] + X_i' \beta_{n,E_i} + \varepsilon_{it}. \tag{11}$$

The $\delta_{h,n}$ coefficients capture the dynamic effects of unionization for elections with $H_i = h$. We also estimate all subsequent treatment-effect heterogeneity using this specification.

Table 2 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 industry. The only marginally significant pre-period estimate is for the service sector, where we find the smallest main effects. Second, the decreases in overall employment and survival are driven by large effects for elections in 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 (SE 0.029) and -0.057 (SE 0.024), respectively. Moreover, the ten-year survival estimate for the service sector is not significantly different than zero, and the confidence interval allows us to reject long-run survival declines of more than four pct. pts.

Appendix Table A3 shows that the smaller effects of unionization in the service sector are robust to alternative controls and sample selection criteria. Specifically, it presents the point estimate

[32] Charles et al. (2023) suggest that the rise in Chinese imports (i.e., the "China Shock) may have increased firms' labor demand elasticities, which adversely affected unions' bargaining position.

[33] We use the Fort and Klimek (2016) harmonized 2012 NAICS codes. We define manufacturing as NAICS sectors 31–33, services as NAICS 51–81 and 44–45, and the other industries as the residual. There is no standard "service sector" definition (e.g., the BLS and Census use different definitions). Compared to the sampling frame for the Census's Service Annual Survey, we include retail trade in the services group and exclude utilities, transportation, and warehousing. This follows Bronfenbrenner (2002) who also excludes utilities, transportation, and warehousing from her classification of service-sector unions. We also include retail because unionized retail workers (e.g., grocery stores) are commonly referred to as part of "service-sector" unionization.

[34] For all heterogeneity analyses, we add the heterogeneity group as a control to account for group-specific trends.

and standard error of the *difference* between the manufacturing and service-sector coefficients at the five- and ten-year time horizons. The effects in manufacturing remain significantly larger when only including pooled industry and employment controls, using a 30–70% vote-share bandwidth, and restricting the sample to elections where the bargaining unit included at least 25% of establishment employment. The last restriction shows that the smaller effects in the service sector are not because service-sector elections are more likely to have a very small share of workers in the bargaining unit.

5.5 Additional Robustness Checks

DiD Weighting Robustness We next show that our results are robust to alternative methods of addressing heterogeneous treatment effect weighting issues in DiD settings. The most prominent issue concerns *cohort* heterogeneity. Our "stacked regression" approach ensures non-negative weights on each cohort-specific treatment effect. However, with *covariate*-specific heterogeneity and baseline controls, our strategy may not estimate the average treatment effect on the treated (ATT) of these heterogeneous effects (see Roth et al. (2022) and the references therein). To address this issue, we re-estimate our main results using Borusyak et al. (2022)'s imputation-style estimator.[35] Appendix Table A4 presents the results for DHS employment and log employment. For the pooled and manufacturing samples, the estimates are qualitatively the same as our baseline results.[36] This similarity confirms that heterogeneous treatment effect weighting issues do change our main results.

Bargaining Unit Size Next, we show that the negative impact of unionization is larger when the bargaining unit includes a greater share of workers (Lee and Mas (2012) conducted a similar test). The motivation is that the relative size of the bargaining unit should mediate the effects of unionization. However, potential violations of our identifying assumption may not be mediated by the share of unionized workers (e.g., workers voting based on expectations of company performance). Appendix Table A5 presents estimates from interacting the treatment indicators with the share of each establishment's employment in the bargaining unit (see Appendix D). For both outcomes, the three- and five-year treatment effects are significantly increasing in the bargaining unit share.[37]

10-Year Pre- and Post-Period Estimates Appendix Figure A5 plots DHS employment growth rate estimates with *ten-year* pre- and post-periods. First, it shows no evidence of large pre-trends in employment growth rates up to ten years before elections for the pooled or manufacturing samples. Although for all industries, we find significant estimates at the six-, seven-, and eight-year horizon,

[35] We implement this procedure using the `did_imputation` package. The only modification we make is to omit the event time $n = -1$ coefficient rather than the earliest period. We make this change because DHS employment growth rates are mechanically missing for $n = -1$, and it makes the pre-trend estimates more comparable to the results from our main specification.

[36] Across all estimates, the largest absolute difference between coefficients is -.108 versus -.07 for manufacturing, five-year, log employment. The point estimates for DHS pre-trends are identical between our stacked TWFE and the imputation-style estimator. Since the DHS growth rate takes long differences relative to event-time $n - 1$ and these specifications do not include establishment-level fixed effects, the pre-trend estimates are already equivalent to only estimating the pre-trend coefficients using untreated observations, like the imputation estimator.

[37] The interactions, however, are not significantly different than zero at the ten-year horizon. One explanation for this is that the bargaining unit share could change substantially in the ten years following the election.

the estimates are economically small (e.g., 1.7 to 2.0% differences). Moreover, the ten-year pre-period estimate is insignificant, and its confidence interval allows us to rule out differences of more than 3.2%. Second, the treatment effects remain stable between three to ten years after the election.

Parametric Vote-Share Tests To complement the previous *nonparametric* vote-share heterogeneity tests, we next estimate *parametric* versions of the same tests. We first test for a linear trend in pre-election employment growth rates across the whole vote-share distribution. Second, we test for linear trends in post-election outcomes separately for winning and losing elections. There are two motivations for these parametric tests. First, these tests may have more power. Second, they provide a parsimonious way to assess robustness to different controls. We show that our estimates are qualitatively the same with the *employment and industry* controls or the *flexible* controls.

To implement these tests, we modify the specification in equation 8. Specifically, instead of interacting event-time with the winning indicator, $\mathbb{1}[V_i > .5]$, we include the following interactions[38]

$$\underbrace{\mathbb{1}[t - E_i = n]}_{\text{Event Time Indicators}} \times \begin{cases} \rho \cdot V_i & \text{if } n < 0 \\ \eta \cdot \mathbb{1}[V_i > .5] + \theta \cdot V_i + \tau \cdot V_i \times \mathbb{1}[V_i > .5] & \text{if } n \geq 0. \end{cases} \tag{12}$$

For the pre-period, ρ estimates a linear trend by vote share. The post-period specification is analogous to a linear RD specification. Consequently, η captures the treatment effect for close elections (i.e., the RD estimate), θ estimates a linear trend in post-election outcomes by vote share for losing elections, and $\theta + \tau$ captures this trend for winning elections.

Appendix Table A6 includes one- to five-year pre-election vote share trends for the DHS employment growth rate (e.g., the ρ estimates). It includes 20–80% vote-share elections. We present separate estimates for all industries and manufacturing and for our two control specifications. Across all estimates, we never find significant pre-election growth rate trends. This complements the previous nonparametric estimates by showing that the lack of pre-trends across the vote-share distribution holds when formally testing for linear trends and with more limited controls.

Appendix Table A7 presents estimates of post-election vote-share trends. We present separate trend estimates for losing elections (i.e., θ) and winning elections (i.e., $\theta + \tau$). This table includes the *flexible control* specification, but Appendix Table A8 shows qualitatively similar results with only the *employment and industry* controls. Motivated by the issues in Figure 6 with exactly 50% elections, we present estimates with and without these elections.

The results for all industries in Table A7 Panel A show significant negative trends in post-election DHS employment growth rates for both losing and winning elections. However, mirroring the nonparametric analysis, when we exclude the 50% elections, we still detect significant trends for winning elections but not for losing elections. For example, we estimate a five-year trend of -0.066 (SE 0.122) for losing elections and -0.389 (SE 0.149) for winning elections. The significant negative estimate for winning elections indicates larger treatment effects for larger margin-of-support

[38] To estimate this specification, we recenter V_i around zero (i.e., subtract .5 from V_i and the indicator variables in Equation 12 are actually $\mathbb{1}[V_i > 0]$). This is analogous to recentering the running variable in RD settings with a non-zero threshold.

elections. The manufacturing estimates in Panel B are similar to the all-industry estimates. Without excluding the 50% elections, we find negative, although insignificant, trends for losing elections. However, dropping the 50% elections results in smaller trends for losing elections and large, although insignificant, vote-share trends for winning elections (e.g., five-year estimates of -0.072 (SE 0.199) for losing elections and -0.406 (SE 0.299) for winning elections). For establishment survival, we never find significant trends for winning or losing elections for either sample.

Overall, the results in Table A7 bolster the nonparametric tests in Figures 5 and 6. First, they show that the lack of post-election trends in employment growth and survival for losing elections holds true when formally testing for linear trends and with more limited controls (although we still drop 50% elections for employment growth). Additionally, the trend estimates for winning elections provide evidence of treatment effect heterogeneity by vote share. Specifically, for the overall DHS employment growth rate estimates, we find significant vote-share heterogeneity.[39] For survival, however, we cannot reject that there is no vote share heterogeneity.

6 Results: Union Avoidance and Employer Opposition

After documenting the large overall impacts of unionization on employment and survival, we explore two new hypotheses for these effects. Specifically, we test whether firms' ability to avoid working with new unions by reallocating production or their general opposition to unions helps explain the overall negative effects. One motivation for exploring these hypotheses is that the conventional explanations for unionization causing establishment closures have not been strongly supported by prior research. Specifically, this research has found little evidence that successful union elections since the 1980s led to large wage increases or productivity declines that could drive the firms out of business (see Appendix C). Both of our alternative explanations can rationalize large employment and survival declines, even if unionization has a relatively small effect on wages and productivity.

For this analysis, we focus on manufacturing elections for three reasons. First, the ways firms can avoid working with new unions differ across sectors. In manufacturing, a common union avoidance tactic was shifting production from unionized to non-unionized establishments within a firm (Bluestone and Harrison, 1982; Verma, 1985). However, in construction, most firms are single-establishment firms, so they cannot shift production across establishments (Butani et al., 2005). Instead, construction firm owners would avoid new unions by opening a non-unionized firm that did the previous work of the unionized firm (Evans and Lewis, 1989). By focusing on manufacturing, we can use the establishment-to-firm linkages in our data to test for union avoidance via production shifting. Second, manufacturing is the largest sector where we find negative effects.[40] Finally, we only have detailed measures of establishment-level productivity in manufacturing that allow us to estimate effect heterogeneity by baseline productivity.

[39] For manufacturing, the estimates are only significant at the 10% level, although Appendix Table A8 shows that when only including the industry and employment controls, we find significant estimates at the 5% level.

[40] For all NAICS sectors not in "services", 54% of elections are in manufacturing, compared to 18% for transportation and warehousing (the next largest sector). When weighted by eligible votes, 69% are manufacturing.

6.1 Union Avoidance via Production Shifting

Our first hypothesis is that firms avoid working with new unions by shifting production from newly unionized establishments to their other establishments. This idea goes back to at least Ulman (1955), who describes the difficulty in unionizing multi-establishment, or multi-plant, firms because "if these two plants are controlled by the same interests, and one of them is shut down, production may be diverted from the idle plant to the plant remaining in operation." Additionally, the hypothesis is supported by evidence from this period that firms shifted employment and investment from their unionized establishments to non-unionized establishments (Verma, 1985; Kochan et al., 1986a).

Multi-Establishment Firm Heterogeneity Since production shifting is only possible for firms with multiple establishments, we estimate whether unionization's effects are larger for elections at multi-establishment (MU) versus single-establishment (SU) firms. We define MU firms as firms with at least two establishments one year before the election.

Figure 8 separately plots the estimates 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 the x-axis, 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 find significantly larger employment declines for elections at MUs at the three- and ten-year horizons. The estimates for SUs are, however, still negative and significant. For the establishment survival estimates, the differences are even more striking. For all time horizons, the effects are significantly larger for MUs, and no SU estimates are significantly different than zero. For example, the ten-year survival estimates are -0.122 (SE 0.021) versus -0.029 (SE 0.029) for MUs and SUs, respectively.

Appendix Table A9 shows robustness to more limited controls and a 30–70% vote-share bandwidth. It presents estimates and standard errors of the difference between the SU and MU estimates. The estimates are very similar with the "ind + emp ctrls." For the 30–70% bandwidth, we estimate substantially larger survival effects for MU firms, but the larger standard errors only lead to a significant difference at the five-year horizon.

These results show that unionization's effect on survival in manufacturing is driven by establishment closings at MU firms. This evidence is consistent with MU firms responding to unionization by shifting production across establishments.

Direct Employment Shifting Estimates Next, we more directly test whether manufacturing firms avoid new unions by shifting production to other establishments. Specifically, we analyze if a successful election at one establishment increases the employment and survival of the firm's other establishments. While the production-shifting hypothesis predicts positive effects on other establishments, other mechanisms like input-output linkages or financial constraints predict negative effects. An additional prediction of the production-shifting hypothesis is that the effects should be largest at the other establishments where it is easiest to produce the same products as the election establishment. Consequently, we start by only including other manufacturing establishments and then restrict to establishments in the same three-digit NAICS industry as the election establishment.[41]

To construct the sample, we start with all manufacturing elections at MU firms. Next, we take all the firms' *other* manufacturing establishments that existed during the election year and never experienced their own union election.[42] We then calculate these establishments' DHS employment growth rates relative to one year before the election. Finally, we estimate a modified version of our main specification 8 with this sample. The two differences from our main specification are that the relative time and vote-share variables are defined from the *other* establishment's election, and we weight the regression by each establishment's share of its firms' employment. The reason for the weighting is that the sample could include multiple establishments 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 establishments). Finally, we two-way cluster the standard errors by the firm ID from the election year and establishment.[43] See Appendix D.1 for more details.

Figure 9 Panel A plots the employment effects of a successful election on the firms' other establishments. The hollow blue estimates include all of the firm's other manufacturing establishments. For all establishments, there is no evidence of higher employment growth following successful elections. However, it is not surprising that we do not find spillovers when we include *all* other establishments. Specifically, many of these establishments may have produced different products than the election establishment, making it difficult to move production to these establishments.

We find significant employment differences when we restrict to establishments that produced similar products to the election establishments. The solid estimates in Panel A only include the firm's other establishments in the same three-digit industry as the election establishment. Two years after the election, we estimate a relative growth rate increase of 0.043 (SE .019) for establishments at firms with successful versus unsuccessful elections. This effect persists three and four years after the election but becomes insignificant after five years. Appendix Table A10 shows that some of the increased employment growth is due to an increased likelihood of survival.

Figure 9 Panel B splits the elections based on whether the *election establishment* constituted a large share of the firm's total employment. The motivation is that we should have less power to detect spillovers when the election establishment was only a small share of the firm's overall employment. We split up elections based on whether the election establishment had over ten percent of the firm's employment in the same three-digit industry during the election year. The estimates in Panel B show that our spillover estimates are driven by relatively large elections. This is reassuring

[41] LaLonde et al. (1996) 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 establishments, where we also do not find spillovers. We only find evidence of spillovers within the same three-digit industry. Bradley et al. (2017) similarly find that firms shift R&D activity away from newly unionized establishments.

[42] We exclude establishments that ever experienced an election so our "spillover estimates" are not contaminated by direct effects. Yet, this conditioning could bias our estimates. The most plausible mechanism, however, biases us *against* finding positive spillovers. Specifically, assume that successful elections lead to *more* future elections at a firm. Since elections occur at relatively fast-growing establishments and the establishment needs to survive to hold a future election, our sample restriction would drop faster-growing establishments at firms with successful elections. This would downward bias our spillover estimates.

[43] Since establishments can change firm IDs, and firms can experience multiple elections, these are not nested.

because these are the elections where we would expect to be able to detect production shifting.

The magnitude of these employment shifting estimates is economically significant. For the same-industry sample, the DHS employment estimates of around 0.04 are consistent with a two pct. pt. increase in the survival probability of the firms' other establishments. As a benchmark, the direct three-year DHS employment effect of unionization for similar elections is -0.23. While our spillover estimates suggest that a sizeable share of the overall negative effects of unionization may be offset by employment shifting, there are several reasons that we cannot use these estimates to calculate this share directly. First, our spillover estimates are at the establishment level, but we need firm-level estimates to calculate the total share offset by reallocation.[44] Second, we focus on the subset of same-industry establishments where we are most likely to detect spillovers. However, calculating the total share offset requires the *firm-level* employment change estimates (e.g., estimates with all other establishments where we have less power to detect significant spillovers).

Overall, this employment spillover evidence is consistent with firms shifting production away from newly unionized establishments. Additionally, the survival estimates suggest that some of this production shifting occurs via decisions over which establishments to close. Although we do not find significant long-run employment spillovers, this does not necessarily indicate a lack of long-run production shifting. First, given the increased variance of long-run growth rates, we may not have enough power to detect effects. Second, we may not be capturing all margins of production shifting over longer time horizons. For example, our analysis excludes production shifting to *new* establishments or to other countries (see Bluestone and Harrison (1982) and Bronfenbrenner (2000) for evidence of international production shifting after unionization).

6.2 Employer Opposition to Unions

Our second hypothesis is that unionization leads to more adverse effects when the firm is more opposed to unions. One explanation for this hypothesis is that some of the employment and survival declines from unionization may be driven by managers' or owners' anti-union animus rather than direct economic costs of unions (see Section 7 for alternative interpretations).[45] This hypothesis also has a long history. For example, Foulkes (1980) documents that some non-unionized firms were motivated by a philosophical opposition to unions, and Leonard (1992) discusses whether the effects of unionization are due to increased costs or anti-union animus. Additionally, anti-union animus is prominent in U.S. labor law. For example, the NLRB and U.S. courts, including the Supreme Court, frequently consider legal cases that hinge on whether an establishment closure was motivated by economic rationale versus anti-union animus.[46] To test this hypothesis, we estimate treatment effect heterogeneity based on two proxies for firms' opposition to the union.

[44]We conduct an establishment-level analysis for two reasons. First, the establishment longitudinal linkages are higher quality than firm-level linkages (Haltiwanger et al., 2013). Second, we may have more power at the establishment level because we can include baseline establishment controls that predict employment growth.

[45]We cannot distinguish between opposition driven by managers' versus owners' preferences. However, in the U.S., a large share of owners are also managers (Kim et al., 2022).

[46]See Section 2 for legal details. The NLRB often considers cases about whether anti-union animus motivated plant closures (Textile Workers Union v. Darlington Mfg. Co., 1965; Weather Tamer. v. NLRB, 1982).

Unionized Versus Non-Unionized Firm Heterogeneity First, we estimate effects separately for elections at MU firms with and without other unionized establishments. The motivation is evidence that non-unionized firms (e.g., firms with no unionized establishments) were more opposed to unions than were (partially) unionized firms. For example, Freedman (1979) and Kochan et al. (1986b) show that firms with lower unionization rates were more committed to remaining non-union and provide accounts of managers at these firms "vigorously resist[ing] dealing with unions."[47]

To test for heterogeneity by firms' unionization status, we split 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 (see Appendix D). Since the FMCS data start in 1984, we classify firms starting in 1985 and show robustness to starting in 1990. Figure 10 shows DHS employment and survival estimates for elections at unionized versus non-unionized firms. For DHS employment, elections at non-unionized firms lead to larger employment decreases than at unionized firms. These differences are significant at the five- and ten-year horizons. For establishment survival, the differences are small and insignificant at the three- and five-year horizons. However, at the ten-year horizon, the negative survival effect is substantially larger for elections at non-unionized firms (e.g., -0.20 (SE 0.040) versus -0.09 (SE 0.027) for non-unionized versus unionized firms, respectively). These estimates show that the long-run negative effects of unionization are larger at non-unionized firms. This evidence is consistent with these firms being more opposed and resistant to unionization.

Appendix Table A11 shows that the larger effects at non-unionized firms are robust to alternative controls and sample selections. It presents estimates when only including the industry and employment controls, classifying unionized versus non-unionized firms starting in 1990, and using a 30–70% bandwidth. For all three alternative specifications, the estimates are qualitatively the same as our baseline estimates.

Election Delay Time Heterogeneity Our second proxy for employer opposition is delay during the election process. The motivation is that employers use tactics that delay the election 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 delay "steals momentum from a union-organizing drive." Additionally, other tactics employers use to influence elections also cause a delay (e.g., challenging the bargaining unit). Finally, delay time is associated with lower election win rates, consistent with it being a proxy for the intensity of anti-union campaigns (Roomkin and Block, 1981; Ferguson, 2008).

We first define delay time and verify that it is related to election outcomes. We define delay time as the number of days between the date the election petition was filed and the date the election was held (see Appendix D). The average delay in our sample is 62 days, and the 10th and 90th percentiles are 31 and 80 days, respectively. Appendix Figure A6 shows that our delay time measure

[47]One reason unionized firms would respond less aggressively to new unionization attempts is that their other unionized workers could pressure the entire firm to discourage aggressive responses. An example is the failure of GM's "southern strategy" of opening non-union plants in the South due to pressure from the UAW (Nelson, 1996).

is negatively associated with election success rates and positively associated with the probability of any challenged votes (another proxy for the anti-union campaign intensity). These relationships remain significant when we control for other election characteristics.

To analyze if unionization's effects differ by delay time, we estimate treatment effects separately for terciles of the within-year delay time distribution. Figure 11 plots DHS employment and survival estimates for the first and third terciles. Panel A includes all elections, and Panel B includes elections at MU firms. In both figures, the effects of unionization on employment and survival are larger for elections in the top tercile. For elections at MUs, the top versus bottom tercile estimates are significantly different for both outcomes at the three- and ten-year horizons (e.g., the 10-year survival effects for the top and bottom terciles are -0.196 (SE 0.037) and -0.071 (SE 0.036), respectively). This evidence is consistent with unionization causing more negative effects when employers campaigned more aggressively against the union. Overall, these results support our hypothesis that employers' opposition to unions plays a role in the overall negative effects of unionization.

Next, we show the results are qualitatively similar using a continuous measure of delay time. This robustness shows that the previous results hold across the entire delay time distribution, including the second tercile. Specifically, we add an interaction between the event-time treatment indicators and *log* delay time to equation 8.[48] Table 3 presents the coefficients on the log delay time interaction for post-election outcomes. The first two columns show that the negative effects of unionization are significantly larger for elections with longer delays across all time horizons. At the 10-year horizon, an approximately 10% increase in delay time is associated with a .7 pct. pt. larger decrease in survival probability.[49] Columns 3 and 4 show robustness to only including the employment and industry controls. Columns 5 and 6 address the concern that elections with longer delay times are simply elections with larger bargaining units. Specifically, we first residualize log delay time on bargaining unit size deciles and interact this residualized measure with event time. The estimates using residualized delay show qualitatively similar results, although the 10-year estimates are only significant at the 10% level.

6.3 Baseline Productivity Heterogeneity Estimates

Finally, we estimate whether unionization leads to larger survival decreases at establishments with lower baseline productivity. If the survival declines that we document are due to the conventional wage or productivity explanations, many theories of firm dynamics would predict larger declines for lower-productivity establishments.[50] Consequently, larger survival declines at less productive establishments would be consistent with the wage or productivity explanations. Alternatively, our

[48]We also control for log delay time interacted with event time to capture its direct effect.

[49]The magnitudes of the continuous and tercile specification estimates are similar. 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.

[50]In Hopenhayn (1992), firms exit when productivity falls below a threshold, implying that lower-productivity firms are more likely to exit after productivity decreases. Empirical research finds that minimum wages and trade liberalization disproportionately decrease survival for less productive firms (Dustmann et al., 2022; Pavcnik, 2002).

hypotheses predict that the negative effects will be the largest at establishments with the most opposition to or ability to avoid unionization, which may be unrelated to baseline productivity.

To measure establishment TFP, we use Cunningham et al. (2022)'s cost-share measures calculated from the Annual Survey of Manufactures and Census of Manufactures. We use within-industry TFP comparisons to address measurement differences across industries. Specifically, we assign each establishment a productivity tercile based on its pre-election, within year, and six-digit industry TFP ranking (see Appendix D). Appendix Figure A7 Panel A plots the heterogeneity estimates for the first and third TFP terciles. The three- and five-year employment and survival effects are larger for lower-productivity establishments. However, these differences are never significant and, at the five-year and ten-year horizon, are not economically large (e.g., -0.066 (SE 0.023) versus -0.041 (SE 0.022) after five years). Panel B shows that these patterns hold when we restrict the sample to MU firms.

Overall, this evidence does not show larger survival effects at less productive establishments. Thus, it is more consistent with our explanations for the negative effects of unionization than the conventional wage and productivity explanations.

7 Discussion

7.1 Employment and Survival Effects of Unionization

Our first contribution is showing that successful union elections substantially decrease establishment employment and survival, especially in manufacturing and other blue-collar and industrial sectors. Relative to past research, our novel empirical strategy avoids biases from only comparing close elections, and we show that the negative effects extend beyond elections just around the 50% threshold. The most comparable results to ours are Frandsen (2021)'s RD estimates. 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 versus 8–10 pct. pts.). This difference could be due to different samples or empirical strategies.[51] Additionally, our smaller employment and survival effects for service-sector elections match Sojourner et al. (2015)'s estimates for nursing home elections. However, even for close elections, our estimates are inconsistent with DiNardo and Lee (2004)'s null survival and employment effects. A potential explanation is that the LBD longitudinal linkages we use to define survival are of higher quality than those in the InfoUSA or LRD data used by DiNardo and Lee (2004).[52] Finally, we provide the first evidence that the effects of unionization on establishment survival vary substantially across sectors.

[51] Frandsen (2021) restricts his sample to elections with at least 20 voters while we only require 6. Since we find larger effects of elections with a higher ratio of eligible voters to employees, this could explain the differences.

[52] See, Jarmin and Miranda (2002) and Crane and Decker (2019) for comparisons of these datasets.

7.2 Implications of Union Avoidance and Employer Opposition

The conventional model of how unionization decreases employment and survival operates through an establishment-level supply and demand framework. First, unions could decrease employment by raising wages and moving the establishment up its labor demand curve (Nickell and Andrews, 1983). Alternatively, unions could lower productivity, which would lower employment by shifting the establishment's labor demand curve (Brown and Medoff, 1978). In this framework, large employment decreases require large establishment-level wage increases, productivity declines, or very elastic labor demand, which the existing research has found little evidence to support. Consequently, our estimates present a puzzle for this framework. Additionally, the union avoidance and employer opposition hypotheses advanced in this paper are both mechanisms absent from this framework. These new mechanisms can explain why unionization could substantially decrease survival and employment even without large establishment-level wage or productivity effects.

First, we show that some of unionization's adverse effects are from firms reallocating production away from newly unionized establishments. This evidence suggests that in manufacturing, some of the establishment-level employment declines represent job *reallocation* rather than job *destruction*. Additionally, although we focus on production shifting in manufacturing, firms in other sectors had different ways of avoiding new unions without shutting down entirely. For example, Hatton (2014) documents firms replacing unionized workers with independent contractors, and Evans and Lewis (1989) document construction firm owners opening separate non-union firms to avoid dealing with a new union (i.e., "Double Breasting"). These examples show that unionization could also lead to employment reallocation in other sectors. Finally, this mechanism can resolve the puzzle discussed above with the standard unionization framework. Specifically, small wage or productivity effects could lead to large establishment survival decreases if firms can cheaply shift production across establishments. We are unaware of estimates of how costly such shifting is, but there is evidence that firms designed their production networks to minimize the cost of reallocating production across plants in response to unionization.[53]

Second, we show that the negative effects of unionization in manufacturing were largest when the firms were likely the most opposed to the union. One interpretation of this result is that the manager's or owners' opposition was driven by their dislike of working with unions rather than the economic costs of unions. This interpretation is consistent with Foulkes (1980), who documents some firms' philosophical opposition to unions. Additionally, it is consistent with Bronfenbrenner (2001), who 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."[54] This interpretation also resolves the previous puzzle because idiosyncratic dislike rather

[53] Bluestone and Harrison (1982) describe how companies "created essentially duplicate production facilities for the same components [...]. The compensation to the company is that a strike or other form of disruption at the original shop can be met by redirecting more production to the non-union facility." This strategy of "parallel production" was commonly used, including by General Motors, General Electric, and Ford.

[54] Another piece of evidence consistent with this interpretation is survey evidence showing that the firms most opposed to unions were not those who expected unions to be the costliest. Freedman (1979) finds that non-unionized firms placed the most weight on resisting unions but expected unions to be least able to bargain for higher wages.

than economic factors could cause the negative employment and survival effects.

An intermediate interpretation is that our employer opposition evidence may reflect hostile labor relations or efforts to prevent the spread of unionization across the firm. For example, our delay-time proxy could reflect unions and management disagreeing more before the election. More adversarial labor relations could drive the negative effects of unionization (Krueger and Mas, 2004; Alder et al., 2023). Alternatively, our finding that entirely non-unionized firms are more likely to close a newly unionized establishment could reflect efforts to prevent the union from spreading across the firm. Specifically, similar to Selten (1978)'s "chain store paradox," a non-unionized firm might close a newly unionized establishment to convey an aggressive stance on unions and deter future elections, even if it would not be economically rational to close that establishment when considering it in isolation.[55] Both of these interpretations are also absent from the standard framework, and they imply that the effects of unionization on employment and survival may substantially overstate the establishment-level economic costs.

On the other hand, it is possible that our employer opposition evidence reflects rational expectations of the direct economic costs of unions (i.e., that the firms most opposed to unions were those where unions would be the costliest). This interpretation is supported by evidence showing direct costs of unions (e.g., negative stock-price and profit effects (Lee and Mas, 2012; Freeman and Medoff, 1984)). However, part of Lee and Mas (2012)'s results does not support the interpretation that the negative effects we document are solely driven by unions' direct costs. In particular, for close union elections, Lee and Mas (2012) do not find stock price declines, whereas we still find negative effects for these elections (Frandsen (2021) also notes this puzzle). One way to reconcile these findings is that the negative employment estimates for close elections are driven by our alternative hypotheses, while economic costs play a larger role for larger margin-of-victory elections. Overall, since we do not estimate unions' direct costs, we cannot rule out the rational cost explanation. Yet, our evidence shows that understanding the cause of employers' opposition to unions is crucial for fully understanding their effects.

7.3 Policy Implications and Areas for Future Research

The role that production shifting and firms' opposition to unions play in U.S. labor relations may be due to the U.S.'s unique establishment-level collective bargaining framework. First, firms are only able to shift production away from unionized establishments because unionization in the U.S. generally occurs at the establishment level rather than at the firm or sector level. Second, the fact that one establishment may be the only unionized establishment within a firm or labor market may exacerbate employers' incentives to oppose unions. For example, firms have a strong incentive to oppose the first union campaign at one of their establishments to prevent unionization from spreading across the firm. Consequently, our analysis suggests that increases in collective bargaining at higher levels (e.g., the firm or sector) may result in more muted negative effects than we document.

[55] For example, Walmart switched to pre-packaged meat in *all* stores after meat cutters at one store unionized in 2000 (Zimmerman, 2000).

Our estimates of effect heterogeneity across sectors add to recent research that finds similar heterogeneity in the impact of other policies that attempt to increase wages. For example, Cengiz et al. (2019) and Harasztosi and Lindner (2019) find negligible employment effects of minimum wages in the service sector but larger negative effects in manufacturing. Since shifting production is often easier for manufacturing firms, the employment reallocation channel we document may help explain why policies to boost wages have more negative effects in manufacturing. Consequently, the optimal policy for raising wages may differ across sectors depending on whether firms can reallocate production to avoid higher wages.[56] Overall, this suggests that lessons from how tax competition between areas leads to a "race to the bottom" in tax rates also apply to policies aimed at raising wages (Giroud and Rauh, 2019; Mast, 2020; Guo, 2023).

Firms in the U.S. vary substantially in the degree to which they oppose unions.[57] The results in this paper suggest that a fruitful area of future research is better understanding why managers and owners at some firms strongly oppose unions, especially since more intense opposition could directly lead to more negative effects of unions. Consequently, it is important to understand whether firms' beliefs about the effects of being unionized may be biased (e.g., overestimating the productivity costs of unions or the likelihood of strikes). This research would complement the literature on workers' and firms' biased beliefs in several other contexts (Heidhues and Kőszegi, 2018; Babcock et al., 2012) and lab experiments showing that individuals have a preference for authority and power in workplace settings, even when not profit maximizing (Fehr et al., 2013; Bartling et al., 2013). Additionally, if some managers or owners have biased beliefs about the impact of unionization on their firms, it would be useful to explore what drives this bias and where they are learned.

8 Conclusion

This paper revisits the effects of successful NLRB union elections on establishment employment and survival. We first introduce a novel research design that extends standard difference-in-differences techniques with falsification tests from the regression discontinuity extrapolation literature. This allows us to avoid biases from vote-share manipulation around the 50% threshold and to estimate treatment effects that include larger margin-of-victory elections. Our strategy and identifying assumption tests can be applied in other panel-data settings where the "forcing variable" is observed. Using this strategy, we show that unionization decreases establishments' employment and their likelihood of survival, particularly in manufacturing and other blue-collar and industrial sectors.

While one interpretation of these negative effects of unionization is that unions lead to large direct costs, we explore two alternative explanations. First, we hypothesize that firms avoid working with new unions by shifting production from newly unionized establishments to other establishments. We support this by showing that the largest effects are at multi-establishment firms and by providing

[56] Mian and Sufi (2014)'s measure of tradable industries does not necessarily identify which industries can shift production across regions. For example, hotels and mines are tradable products but are relatively immobile.

[57] Recently, Amazon adopted an aggressive anti-union stance, while Microsoft agreed to remain neutral in union campaigns (Streitfeld, 2021; Scheiber, 2023)

evidence of increased employment at firms' *other* establishments following successful elections. Second, we hypothesize that unionization leads to more adverse effects when the firm is more opposed to the union. Supporting this, we find the largest effects for elections at non-unionized firms and for elections with the longest delay during the election process, both proxies for employers' opposition to the union. This evidence supports our new hypotheses for why unionization decreases establishment employment and survival, even without direct wage or productivity effects. Overall, this shows that efforts to increase collective bargaining in the U.S. should also address employers' ability to avoid unions and their overall opposition to unions in order to mitigate the negative employment effects we document.

References

Alder, Simeon, David Lagakos, and Lee E Ohanian (2023) "Labor Market Conflict and the Decline of the Rust Belt," *Journal of Political Economy*, Vol. 131.

Angrist, Joshua D. and Miikka Rokkanen (2015) "Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away From the Cutoff," *Journal of the American Statistical Association*, Vol. 110, pp. 1331–1344.

Babcock, Linda, William J Congdon, Lawrence F Katz, and Sendhil Mullainathan (2012) "Notes on behavioral economics and labor market policy," *IZA Journal of Labor Policy*, Vol. 1, p. 2.

Bartling, Björn, Ernst Fehr, and Klaus M. Schmidt (2013) "Use And Abuse Of Authority: A Behavioral Foundation Of The Employment Relation," *Journal of the European Economic Association*, Vol. 11, pp. 711–742.

Bennett, Magdalena (2020) "How Far is Too Far? Estimation of an Interval for Generalization of a Regression Discontinuity Design Away from the Cutoff," *Working Paper*.

Bluestone, Barry and Bennett Harrison (1982) *The Deindustrialization of America: Plant Closings, Community Abandonment, and the Dismantling of Basic Industry*, New York: Basic Books, Inc., Publishers.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2022) "Revisiting Event Study Designs: Robust and Efficient Estimation," *Working Papers*, arXiv:2108.12419 [econ].

Bradley, Daniel, Incheol Kim, and Xuan Tian (2017) "Do Unions Affect Innovation?" *Management Science*, Vol. 63, pp. 2251–2271.

Bronfenbrenner, Kate (1996) "Final Report: the Effects of Plant Closing or Threat of Plant Closing on the Right of Workers to Organize," *Report Submitted to the The Labor Secretariat of North American Commission for Labor Cooperation*.

——— (1997) "The Role of Union Strategies in NLRB Certification Elections," *ILR Review*, Vol. 50, pp. 195–212.

——— (2000) "Uneasy Terrain: The Impact Of Capital Mobility On Workers, Wages, And Union Organizing," *Report Submitted to the U.S. Trade Deficit Review Commission*.

——— (2001) "Uneasy Terrain: The Impact Of Capital Mobility On Workers, Wages, And Union Organizing Part II: First Contract Supplement," *Report Submitted to the U.S. Trade Deficit Review Commission*.

——— (2002) "Overcoming the Challenges to Organizing in Manufacturing," *Report Submitted to the AFL-CIO*.

Bronfenbrenner, Kate and Tom Juravich (1998) "It Takes More Than House Calls: Organizing to Win with a Comprehensive Union-Building Strategy," in *Organizing to Win: New Research on Union Strategies, Kate Bronfenbrenner and Sheldon Friedman, et al. (eds.)*, Ithaca, NY: ILR Press, pp. 19–36.

Brown, Charles and James Medoff (1978) "Trade Unions in the Production Process," *Journal of Political Economy*, Vol. 86, pp. 355–378.

Butani, Shail, George Werking, and Vinod Kapani (2005) "Employment dynamics of individual companies versus multicorporations," *Monthly Labor Review*, p. 13.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019) "The Effect of Minimum Wages on Low-Wage Jobs," *The Quarterly Journal of Economics*, Vol. 134, pp. 1405–1454.

de Chaisemartin, Clément and Xavier D'Haultfoeuille (2020) "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, Vol. 110, pp. 2964–2996.

Charles, Kerwin Kofi, Matthew S. Johnson, and Nagisa Tadjfar (2023) "Trade Competition and the Decline in Union Organizing: Evidence from Certification Elections," *NBER Working Papers*.

Chodorow-Reich, Gabriel (2014) "The Employment Effects of Credit Market Disruptions: Firm-level Evidence from the 2008-9 Financial Crisis," *The Quarterly Journal of Economics*, Vol. 129, pp. 1–59.

Chow, Melissa, Teresa Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elisabeth Ruth Perlman, Martha Stinson, and T. Kirk White (2021) "Redesigning the Longitudinal Business Database," *NBER Working Paper*, p. w28839, Place: Cambridge, MA.

Crane, Leland D. and Ryan A. Decker (2019) "Business Dynamics in the National Establishment Time Series (NETS)," *Finance and Economics Discussion Series Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board*, Vol. 2019.

CRS (2013) "The National Labor Relations Act (NLRA): Union Representation Procedures and Dispute Resolution," *Congressional Research Service Report*, p. 38.

Cunningham, Cindy, Lucia Foster, Cheryl Grim, John Haltiwanger, Sabrina Wulff Pabilonia, Jay Stewart, and Zoltan Wolf (2022) "Dispersion in Dispersion: Measuring Establishment-Level Differences in Productivity," *Review of Income and Wealth*.

Textile Workers Union v. Darlington Mfg. Co., 380 U.S. 263 (1965) .

Davis, Steven J, John Haltiwanger, Kyle Handley, Josh Lerner, and Javier Miranda (2014) "Private Equity, Jobs, and Productivity," *American Economic Review*, Vol. 104, pp. 3956–3990.

Davis, Steven J., John C. Haltwanger, and Scott Schuh (1996) *Job Creation and Destruction*, Cambridge, MA, USA: MIT Press.

DiNardo, John and David S. Lee (2004) "Economic Impacts of New Unionization on Private Sector Employers: 1984-2001," *Quarterly Journal of Economics*, Vol. 119, pp. 1383–1441.

Dinlersoz, Emin and Jeremy Greenwood (2016) "The rise and fall of unions in the United States," *Journal of Monetary Economics*, Vol. 83, pp. 129–146.

Dinlersoz, Emin, Jeremy Greenwood, and Henry Hyatt (2017) "What Businesses Attract Unions? Unionization over the Life Cycle of U.S. Establishments," *ILR Review*, Vol. 70, pp. 733–766.

Dube, Arindrajit, Ethan Kaplan, and Owen Thompson (2016) "Nurse Unions and Patient Outcomes," *ILR Review*, Vol. 69, pp. 803–833.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson (1989) "The Growth and Failure of U. S. Manufacturing Plants," *The Quarterly Journal of Economics*, Vol. 104, p. 671.

Dustmann, Christian, Attila Lindner, Uta Schönberg, Matthias Umkehrer, and Philipp vom Berge (2022) "Reallocation Effects of the Minimum Wage," *The Quarterly Journal of Economics*, Vol. 137, pp. 267–328.

Evans, Stephen and Roy Lewis (1989) "Union Organisation, Collective Bargaining And The Law: An Anglo-American Comparison Of The Construction Industry," *Comparative Labor Law Journal*, p. 33.

Farber, Henrys (2001) "Union Success in Representation Elections: Why Does Unit Size Matter?" *ILR Review*, Vol. 54, pp. 329–348.

Fehr, Ernst, Holger Herz, and Tom Wilkening (2013) "The Lure of Authority: Motivation and Incentive Effects of Power," *American Economic Review*, Vol. 103, pp. 1325–1359.

Ferguson, John-Paul (2008) "The Eyes Of The Needles: A Sequential Model Of Union Organizing Drives, 1999–2004," *Industrial And Labor Relations Review*, Vol. 62, p. 19.

Fort, Teresa C and Shawn Klimek (2016) "The Effects Of Industry Classification Changes On Us Employment Composition," *CES Working Paper 18-28*, p. 24.

Foulkes, Fred (1980) *Personnel Policies in Large Nonunion Companies*, Englewood Cliffs, N.J.: Prentice-Hall.

Frandsen, Brigham (2021) "The Surprising Impacts of Unionization: Evidence from Matched Employer-Employee Data," *Journal of Labor Economics*, Vol. 39, p. 56.

Frandsen, Brigham R. (2017) "Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete," in Matias D. Cattaneo and Juan Carlos Escanciano eds. *Advances in Econometrics*, Vol. 38: Emerald Publishing Limited, pp. 281–315.

Freedman, Audrey (1979) "Managing Labor Relations," *Conference Board Report*.

Freeman, Richard B and Morris M Kleiner (1990a) "Employer Behavior in the Face of Union Organizing Drives," *ILRR*, Vol. 43, p. 16.

Freeman, Richard B. and Morris M. Kleiner (1990b) "The Impact of New Unionization on Wages and Working Conditions," *Journal of Labor Economics*, Vol. 8, pp. S8–S25.

Freeman, Richard and Morris Kleiner (1999) "Do Unions Make Enterprises Insolvent?" *ILR Review*, Vol. 52.

Freeman, Richard and James Medoff (1984) *What Do Unions Do?*, N.Y.: Basic Books.

GAO (2002) "Collective Bargaining Rights: Information on the Number of Workers with and without Bargaining Rights," *U.S. General Accounting Office Report GAO-02-835*.

Giroud, Xavier and Holger M. Mueller (2017) "Firm Leverage, Consumer Demand, and Employment Losses during the Great Recession," *The Quarterly Journal of Economics*, Vol. 132, pp. 271–316.

Giroud, Xavier and Joshua Rauh (2019) "State Taxation and the Reallocation of Business Activity: Evidence from Establishment-Level Data," *Journal of Political Economy*, Vol. 127, p. 55.

Goodman-Bacon, Andrew (2021) "Difference-in-differences With Variation In Treatment Timing," *Journal of Econometrics*, Vol. 225, p. 49.

Greenhouse, Steven (2011) "Labor Board Tells Boeing New Factory Breaks Law," *The New York Times*, Section: Business.

Guo, Audrey (2023) "The Effects of State Business Taxes on Plant Closures: Evidence from Unemployment Insurance Taxation and Multi-Establishment Firms," *Review of Economics and Statistics*, Vol. 105, pp. 580–595.

Haltiwanger, John, Ron S. Jarmin, and Javier Miranda (2013) "Who Creates Jobs? Small versus Large versus Young," *Review of Economics and Statistics*, Vol. 95, pp. 347–361.

Harasztosi, Peter and Attila Lindner (2019) "Who Pays for the Minimum Wage?" *American Economic Review*, Vol. 109, p. 110.

Hatton, Erin (2014) "Temporary Weapons: Employers' Use of Temps against Organized Labor," *ILR Review*, Vol. 67, pp. 86–110.

Heidhues, Paul and Botond Kőszegi (2018) "Chapter 6 - Behavioral Industrial Organization," in B. Douglas Bernheim, Stefano DellaVigna, and David Laibson eds. *Handbook of Behavioral Economics: Applications and Foundations 1*, Vol. 1 of Handbook of Behavioral Economics - Foundations and Applications 1: North-Holland, pp. 517–612.

Heneman, Herbert G and Marcus H Sandver (1983) "Predicting The Outcome Of Union Certification Elections: A Review Of The Literature," *Industrial And Labor Relations Review*, Vol. 36, p. 24.

Hopenhayn, Hugo A. (1992) "Entry, Exit, and firm Dynamics in Long Run Equilibrium," *Econometrica*, Vol. 60, p. 1127.

Jarmin, Ron S. and Javier Miranda (2002) "The Longitudinal Business Database," *SSRN Electronic Journal*.

Kim, Mee Jung, David Brown, John Earle, Kyung Min Lee, and Jared Wold (2022) "Why are Black-Owned Businesses Smaller?" *Working Paper*.

Kleiner, Morris M. (2001) "Intensity of management resistance: understanding the decline of unionization in the private sector," *Journal of Labor Research*, Vol. 22, pp. 519–540.

Knepper, Matthew (2020) "From the Fringe to the Fore: Labor Unions and Employee Compensation," *The Review of Economics and Statistics*, Vol. 102, pp. 98–112.

Kochan, Thomas A, Katz Harry, and McKersie Robert (1986a) *The Transformation of American Industrial Relations*, New York: Basic Books.

Kochan, Thomas A, Robert McKersie, and John Chalykoff (1986b) "The Effects of Corporate Strategy and Workplace Innovations on Union Representation," *ILR Review*, p. 16.

Krueger, Alan and Alexandre Mas (2004) "Strikes, Scabs, and Tread Separations: Labor Strife and the Production of Defective Bridgestone/Firestone Tires," *The Journal of Political Economy*, Vol. 112.

LaLonde, Robert, Gerard Marschke, and Kenneth Troske (1996) "Using Longitudinal Data on Establishments to Analyze the Effects of Union Organizing Campaigns in the United States," *Annales d'Économie et de Statistique*, p. 155.

Lee, David S. and Alexandre Mas (2012) "Long-Run Impacts of Unions on Firms: New Evidence from Financial Markets, 1961–1999," *The Quarterly Journal of Economics*, Vol. 127, pp. 333–378.

Leonard, Jonathan (1992) "Unions and Employment Growth," in *Labor Market Institutions and the Future Role of Unions Edited by Mario Bognanno and Morris Kleiner*: Blackwell Publishers.

Levitt, Martin and Terry Conrow (1993) *Confessions of a Union Buster*, New York: Crown Publishers.

Logan, John (2002) "Consultants, Lawyers, And The 'union Free' Movement In The USA Since The 1970s," *Industrial Relations Journal*, Vol. 33, pp. 197–214.

Mast, Evan (2020) "Race to the Bottom? Local Tax Break Competition and Business Location," *American Economic Journal: Applied Economics*, Vol. 12, pp. 288–317.

McAlevey, Jane (2020) *A Collective Bargain: Unions, Organizing, and the Fight for Democracy*, New York, NY: Ecco, an imprint of Harper Collins Publishers.

Mian, Atif and Amir Sufi (2014) "What Explains the 2007-2009 Drop in Employment?" *Econometrica*, Vol. 82, pp. 2197–2223.

Munger, Peter, Stephen X Mungertt, and Thomas J Mungerttt (1988) "Plant Closures And Relocations Under The National Labor Relations," *Georgia State University Law Review*, Vol. 5, p. 41.

Nelson, Douglas (1996) "The Political Economy of U.S. Automobile Protection," in *The Political Economy of American Trade Policy*: University of Chicago Press, pp. 133–196.

Nickell, S. J. and M. Andrews (1983) "Unions, Real Wages And Employment In Britain 1951–79," *Oxford Economic Papers*, Vol. 35, pp. 183–206.

Weather Tamer. v. NLRB, 676 F.2d 483 (1982) .

Pavcnik, Nina (2002) "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *The Review of Economic Studies*, Vol. 69, pp. 245–276.

Roomkin, Myron and N Block (1981) "Case Processing Time And The Outcome Of Representation Elections: Some Empirical Evidence," *University Of Illinois Law Review*, p. 25.

Roth, Jonathan and Pedro H. C. Sant'Anna (2023) "When Is Parallel Trends Sensitive to Functional Form?" *Econometrica*, Vol. 91, pp. 737–747.

Roth, Jonathan, Pedro H C Sant'Anna, John Poe, and Alyssa Bilinski (2022) "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *Journal of Econometrics*.

Sandver, Marcus Hart and Kathryn J. Ready (1998) "Trends in and determinants of outcomes in multi-union certification elections," *Journal of Labor Research*, Vol. 19, pp. 165–172.

Scheiber, Noam (2023) "Microsoft Agrees to Remain Neutral in Union Campaigns," *The New York Times*.

Schmitt, John and Ben Zipperer (2009) "Dropping the Ax: Illegal Firings During Union Election Campaigns, 1951-2007," *CEPR Working Papers*, p. 22.

Selten, Reinhard (1978) "The chain store paradox," *Theory and Decision*, Vol. 9, pp. 127–159.

Sojourner, Aaron J., Brigham R. Frandsen, Robert J. Town, David C. Grabowski, and Min M. Chen (2015) "Impacts of Unionization on Quality and Productivity: Regression Discontinuity Evidence from Nursing Homes," *ILR Review*, Vol. 68, pp. 771–806.

Sojourner, Aaron and Jooyoung Yang (2022) "Effects of Union Certification on Workplace-Safety Enforcement: Regression-Discontinuity Evidence," *ILR Review*, Vol. 75, pp. 373–401, Publisher: Sage Publications Inc.

Streitfeld, David (2021) "How Amazon Crushes Unions," *The New York Times*.

Sun, Liyang and Sarah Abraham (2021) "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, Vol. 225, pp. 175–199.

Traxler, Franz (1994) "Collective Bargaining: Levels and Coverage," in *OECD Employment Outlook*: OECD.

Ulman, Lloyd (1955) *The Rise of the National Trade Union*, Cambridge, Massachusetts: Harvard University Press.

Verma, Anil (1985) "Relative Flow of Capital to Union and Nonunion Plants Within a Firm," *Industrial Relations*, Vol. 24, pp. 395–405.

Wamsley, Laurel (2017) "Billionaire Owner Shuts Down DNAinfo, Gothamist Sites A Week After Workers Unionize," *NPR*, Section: America.

Weiler, Paul (1983) "Promises to Keep: Securing Workers' Rights to Self-Organization under the NLRA," *Harvard Law Review*, Vol. 96, p. 1769.

Zimmerman, Anne (2000) "Pro-Union Butchers at Wal-Mart Win a Battle, but Lose the War," *Wall Street Journal*, Section: Front Section.

9 Figures

Figure 1 Note: Figure 1 presents three 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 D for details).

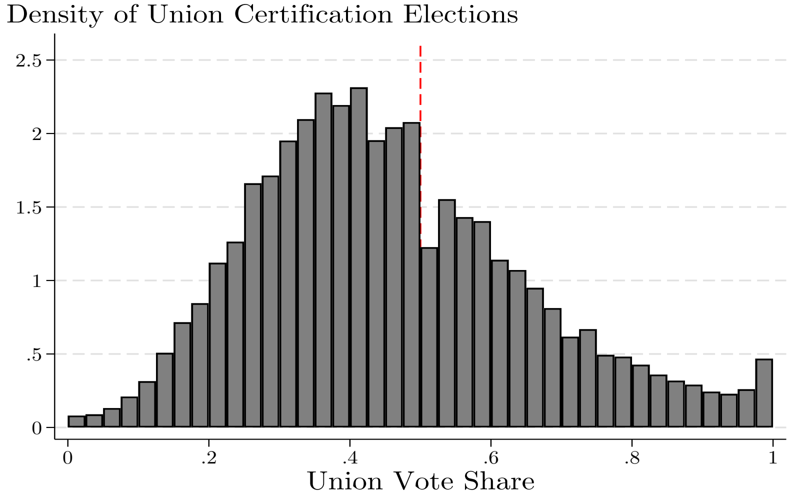
Panel A plots the vote-share histogram of elections with more than 50 total voters. Given the discreteness of the running variable and the fact that our sample includes elections with a small number of votes, it is difficult to detect manipulation from the vote-share density figure for the entire sample, so we restrict the sample to elections with at least 50 votes. See Frandsen (2017) for evidence of manipulation using formal tests that account for the discrete running variables. See Appendix Figure A1 for the vote-share histogram that includes all elections in our sample.

Panel B plots the average and median number of days between the union election date and the date that the case closed. Note, this measure is different than the measure of "delay time" used in Section 6. That measure of delay time is the difference between the date the union files the election petition and the actual election date.

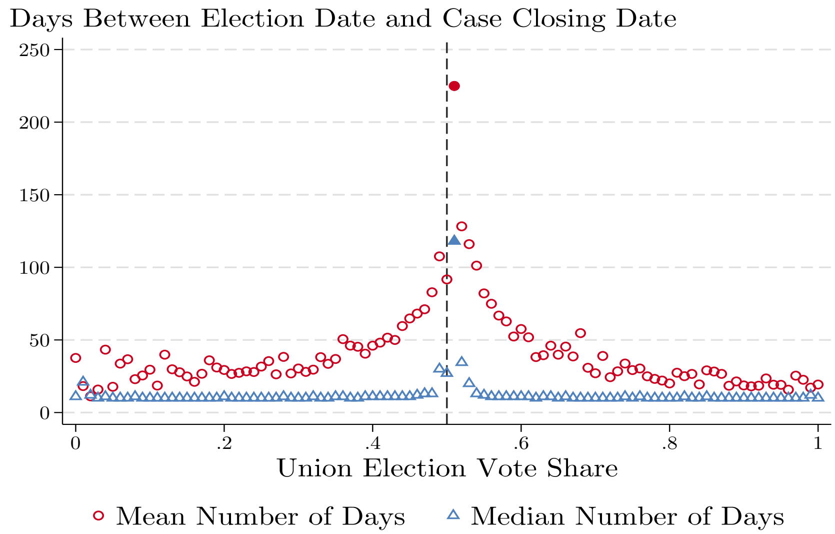
Panel C plots the probability of each union election experiencing a decertification election in the five years following the case closing. The decertification elections are also from our combined NRLB datasets but excluded from our main analysis. For the decertification figure, we match based on exact company names and cities rather than the SoftTFIDF algorithm we use for the main analysis. We estimate that some losing elections (e.g., elections with less than 50% vote shares) have future decertification elections for several reasons. First, a different set of workers at these establishments may have already been unionized (e.g., a different bargaining unit) and later decertified. Second, this could reflect matching errors (e.g., incorrectly matching a future decertification election). Finally, these establishments could have been unionized in a follow-up unionization election.

Figure 1: Characteristics of Close Union Elections

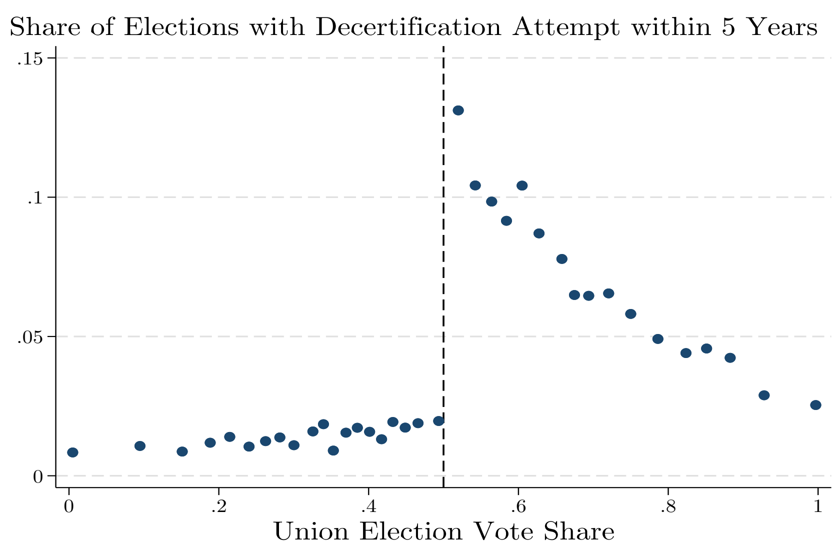
Panel A. Election Vote-Share Histogram, 50 + Vote Elections



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

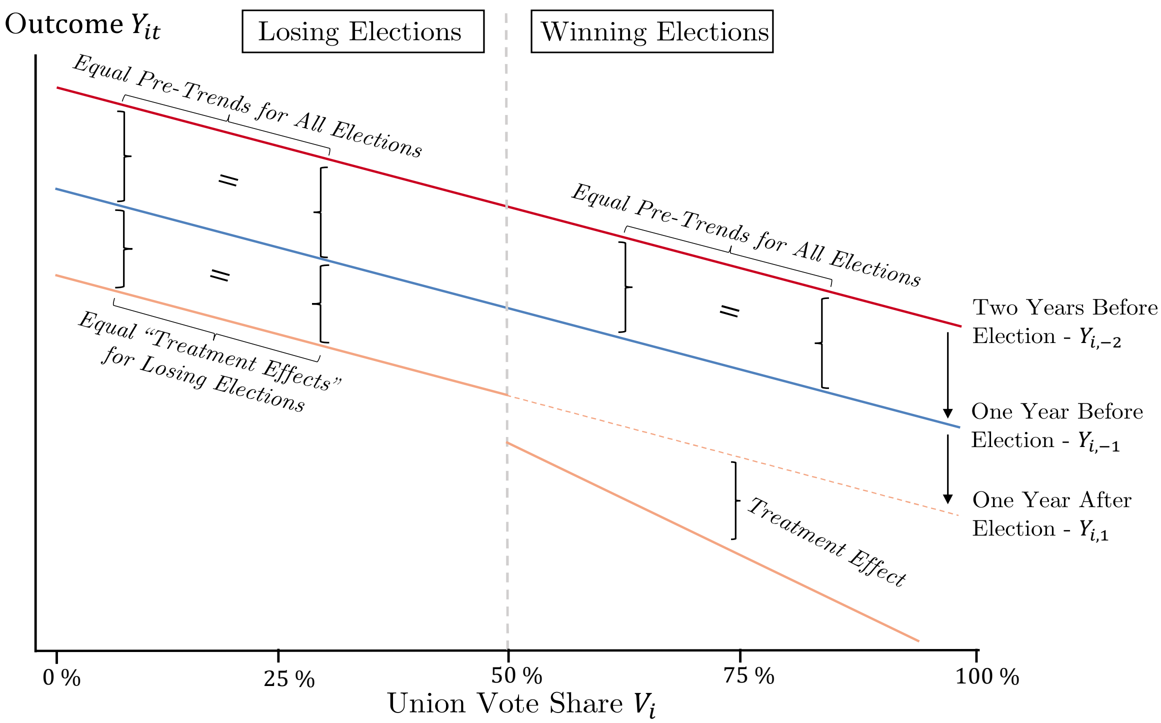


Panel C. Probability of a *Decertification Election* Five Years Following an Election



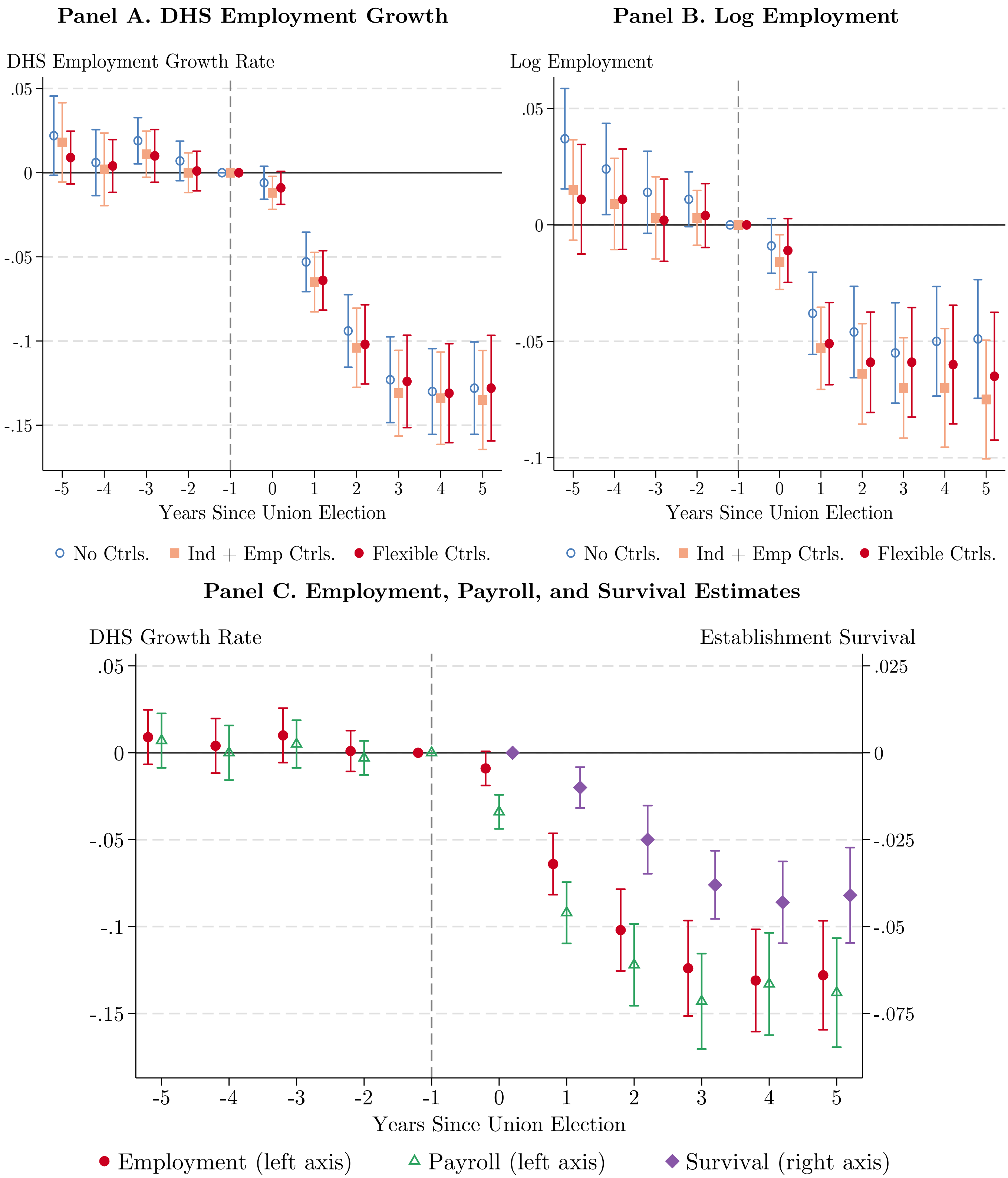
Note: See the previous page.

Figure 2: Testable Implications of Parallel Trends Identifying Assumption



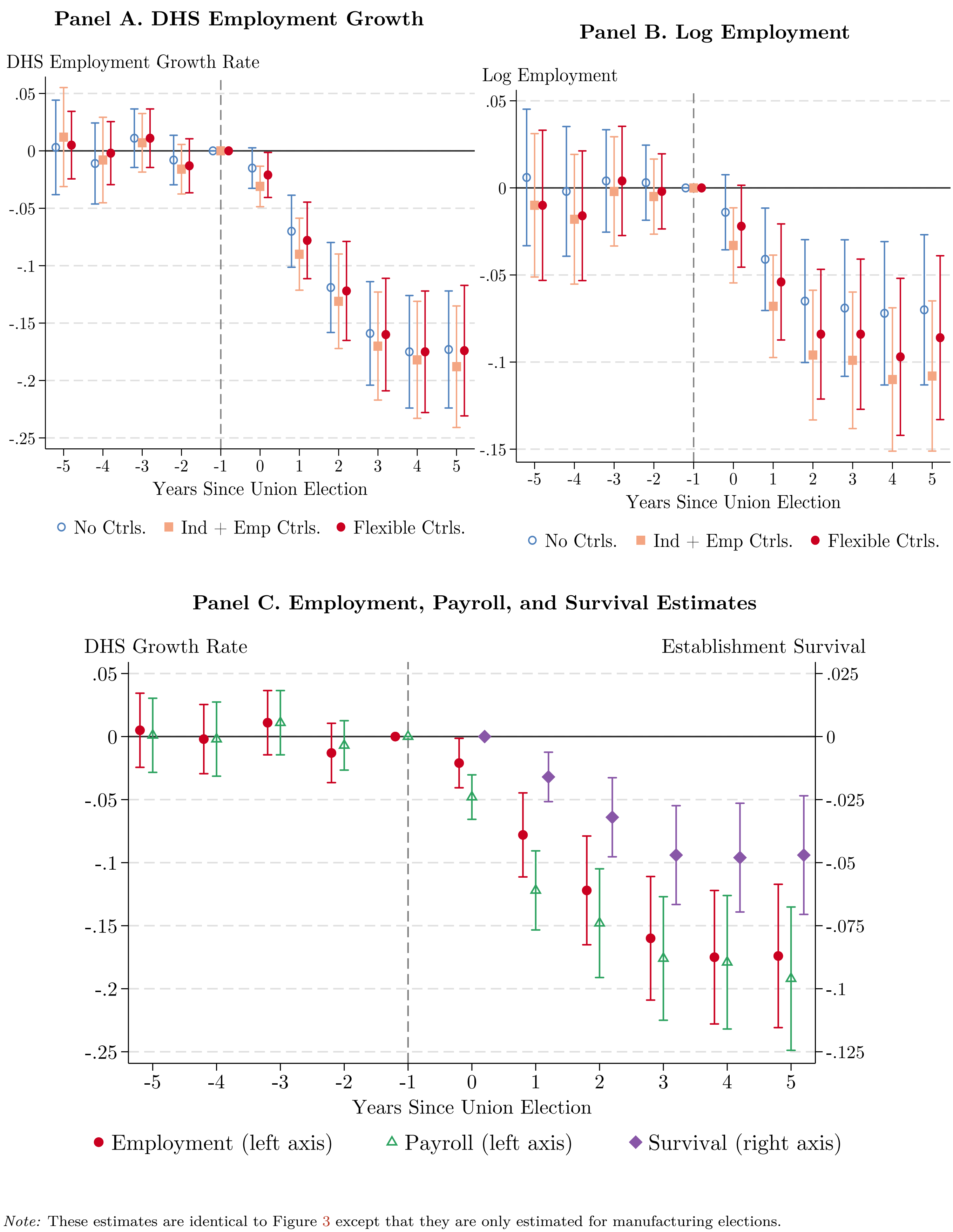
Note: This figure illustrates our empirical strategy's identifying assumption and its testable implications discussed in section 4. It plots hypothetical average establishment-level outcomes before and after union elections with different vote shares. $Y_{i,-2}$ and $Y_{i,-1}$ correspond to outcomes 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 $Y_{i,-1}$ and $Y_{i,1}$ for losing elections.

Figure 3: Employment and Survival Estimates, 20–80% Vote-Share Elections, All Industries



Note: This figure plots the δ_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).

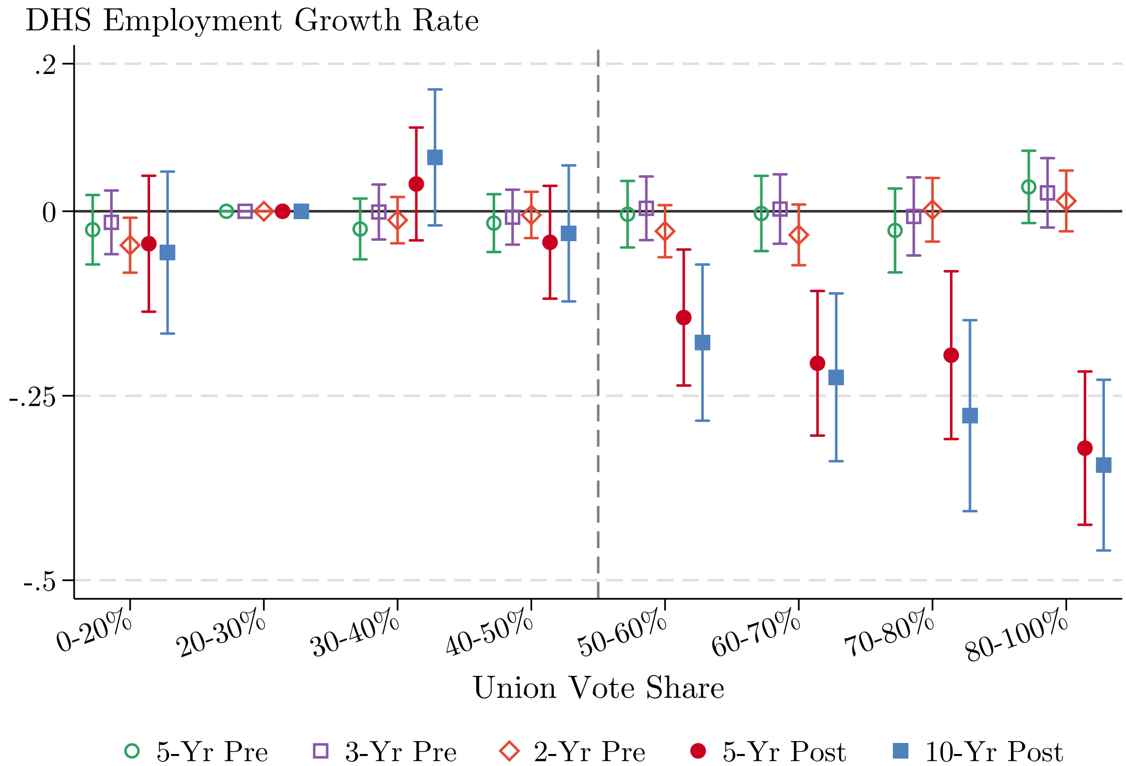
Figure 4: Employment and Survival Estimates, 20–80% Vote-Share Elections, Manufacturing



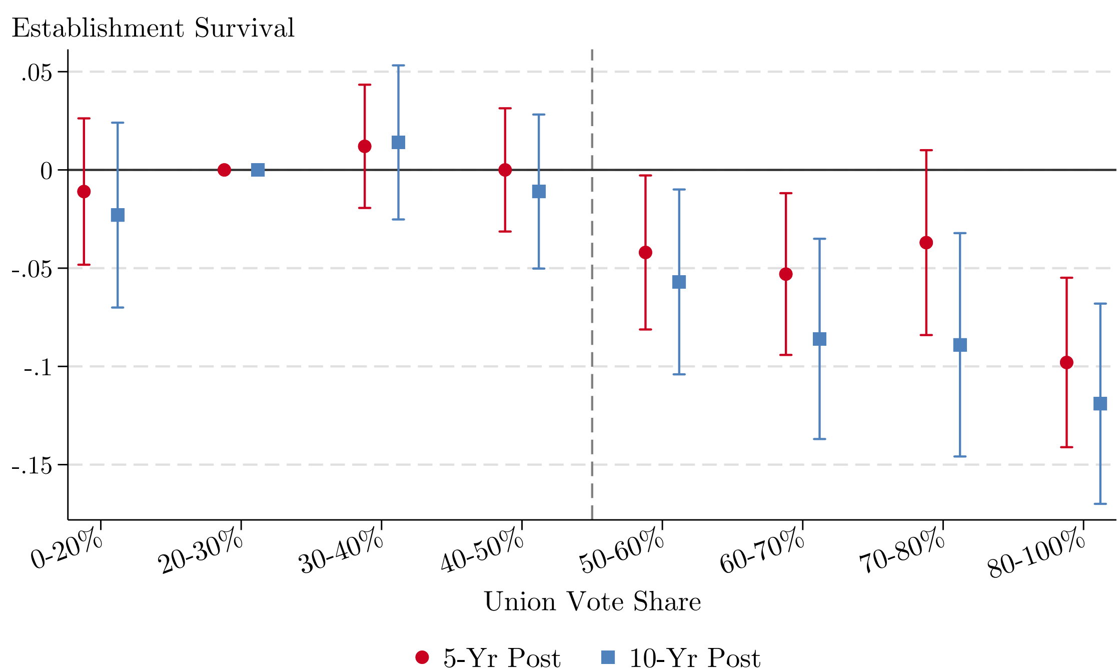
Note: These estimates are identical to Figure 3 except that they are only estimated for manufacturing elections.

Figure 5: Nonparametric Vote-Share Heterogeneity Estimates, Manufacturing

Panel A. DHS Employment Growth Rate



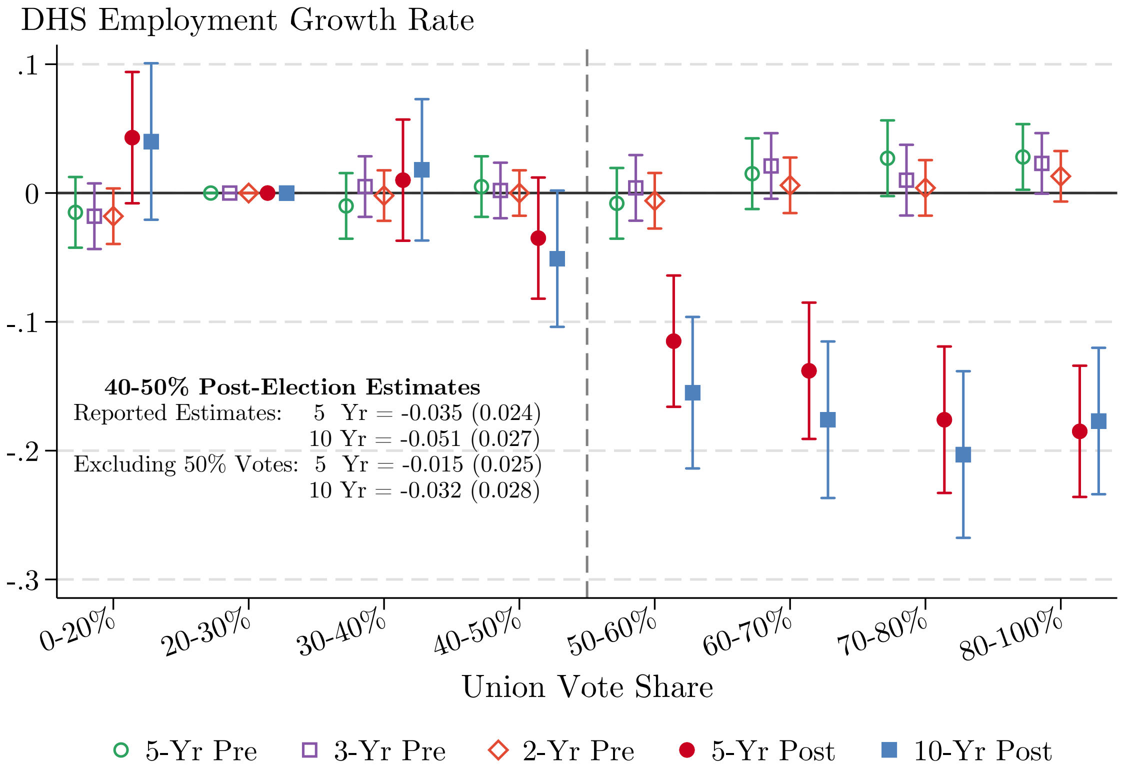
Panel B. Establishment Survival



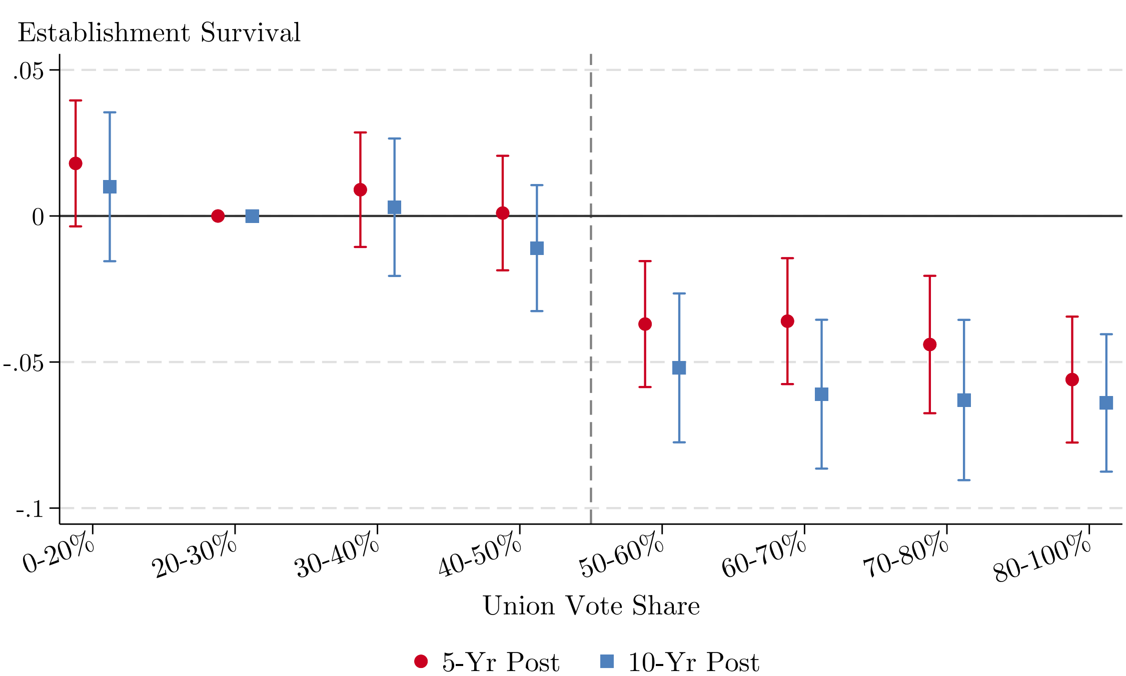
Note: This figure plots the $\delta_{g,n}$ coefficients from estimating the vote-share heterogeneity specification 10 with the vote-share distribution partitioned into eight groups indicated on the x-axis. We omit the 20–30% election group, so the other estimates are relative to that group. The sample 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 of 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

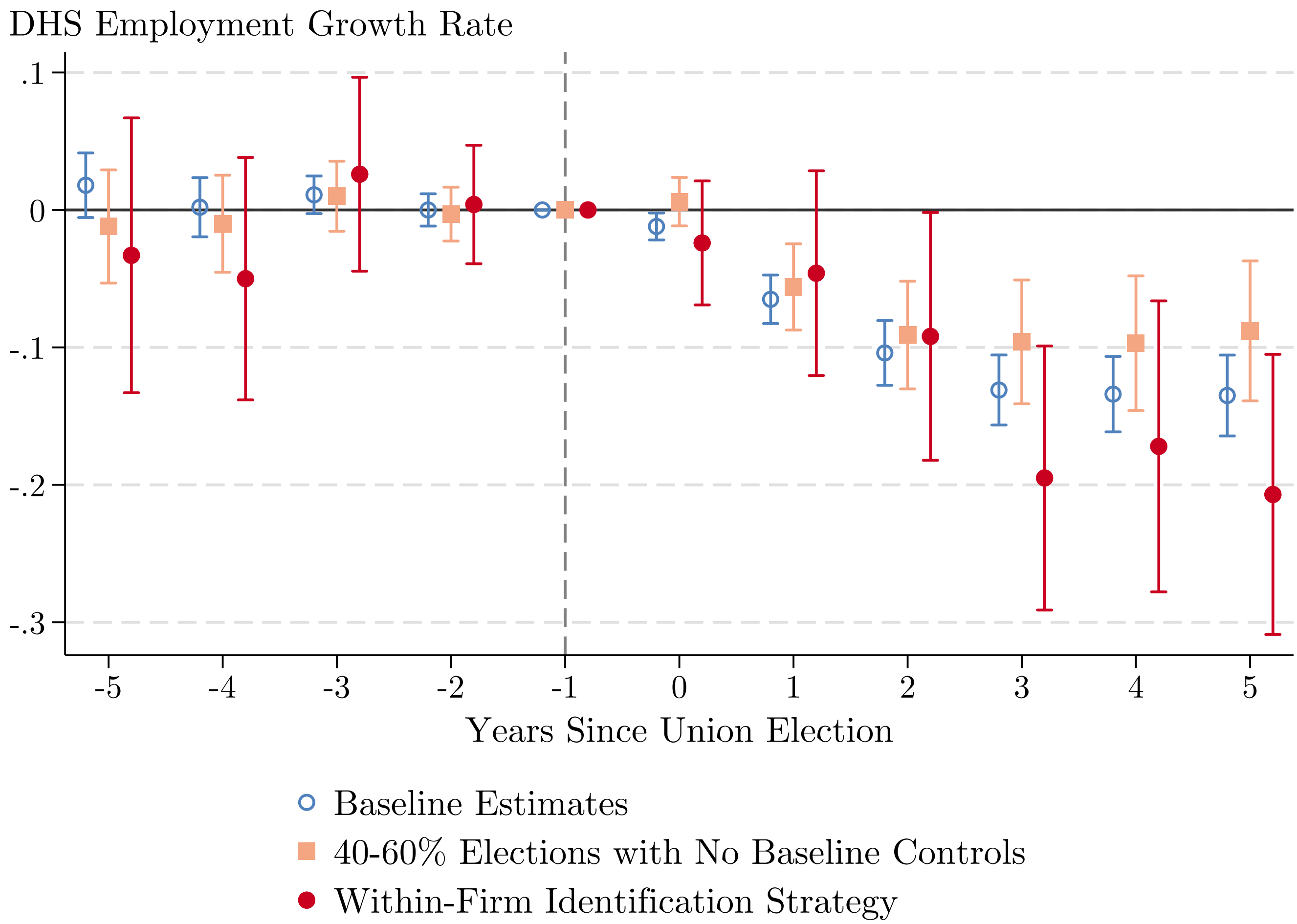


Panel B. Establishment Survival



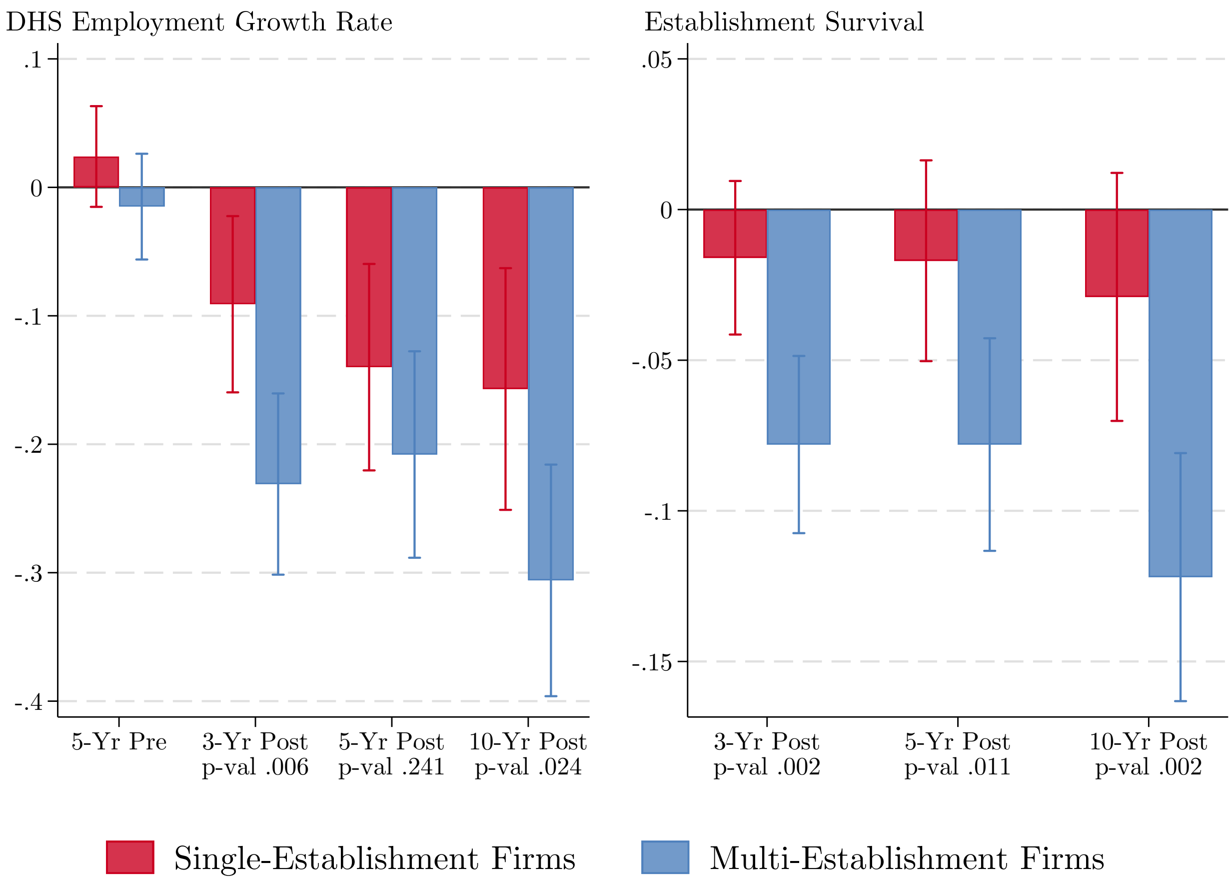
Note: This figure is identical to Figure 5 except it includes elections across all industries. The alternative estimates listed in the text box in Panel A. are the 40–50% estimates excluding elections with exactly 50% of votes (rather than restrict the sample, we include a separate category for 50% vote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

Figure 7: Alternative Identification Strategies, DHS Employment Estimates



Note: This figure plots the δ_n coefficients (i.e., the interaction between winning a union election and being n years from the election) from estimating specification 8 for three different samples and sets of controls. The outcome for all specifications is establishment-level DHS employment growth relative to time -1. The blue points are estimates from our main strategy with employment by industry controls (e.g., the same estimates presented in Figure 3 Panel A). The orange "40-60% Elections" points are from estimating specification 8 only including elections in all industries with vote shares between 40–47.5% and 52.5-60%. For this specification, we only include year-by-cohort fixed effects as controls. The red "Within-Firm" points are from estimating specification 8 only including elections at firms that had at least one winning and losing election in the same year, after imposing all of the other sample restrictions for our main strategy. For this specification, we only include firm-by-year-by-cohort fixed effects. 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).

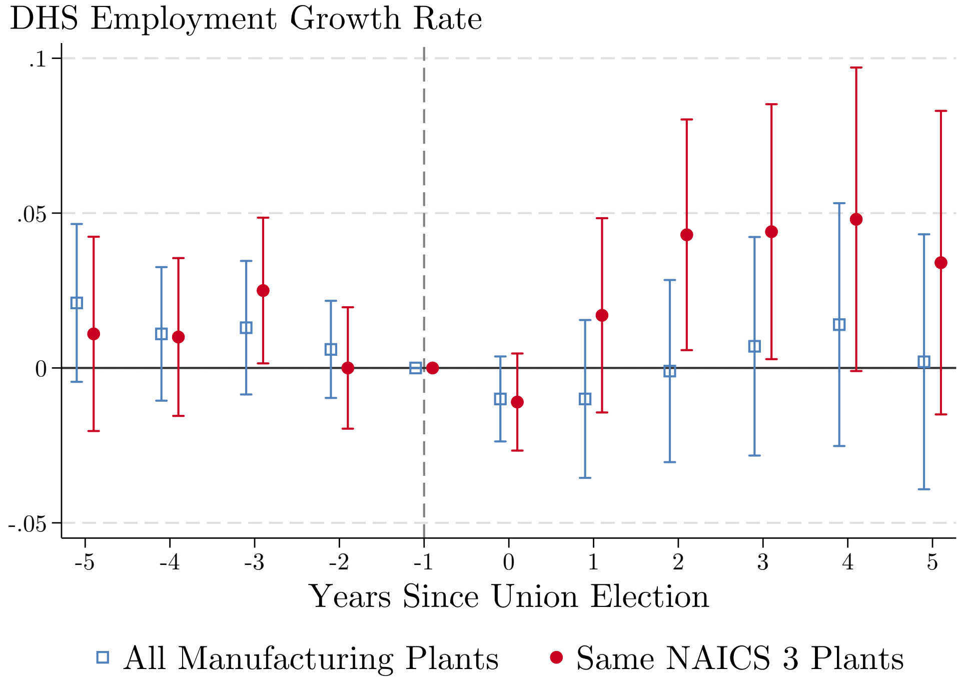
Figure 8: Single- Versus Multi-Establishment Firm Heterogeneity



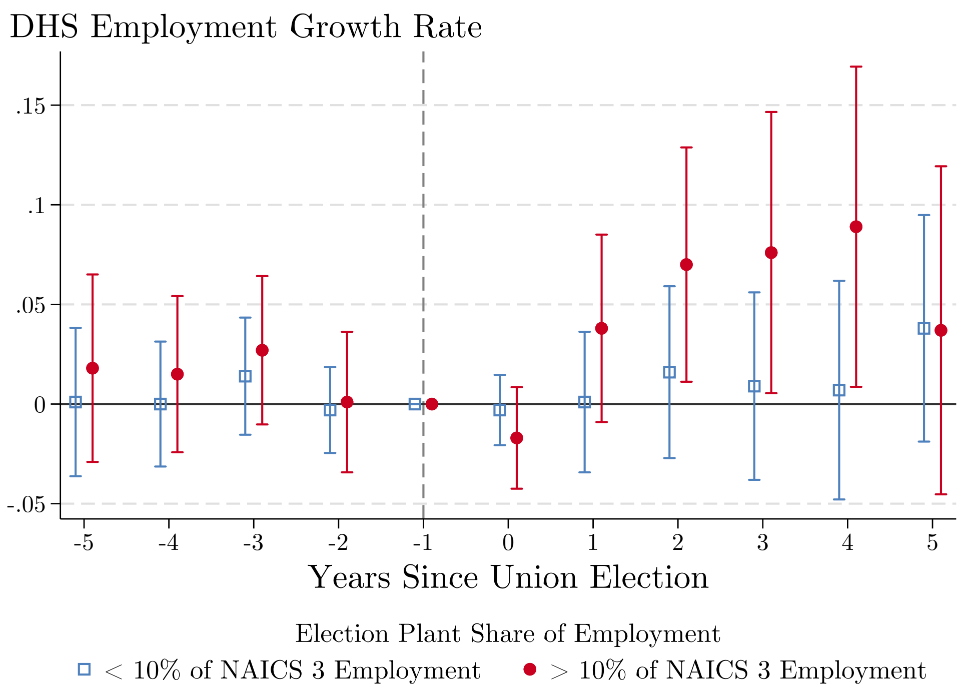
Note: This figure plots the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 11 for elections at single- versus multi-establishment firms. An election at a multi-establishment firm is defined based on whether the firm has any other establishments one year before the election. The sample includes all manufacturing union elections with 20–80% vote shares inclusive. It includes observations from -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 of establishment survival. The estimates include the flexible control specification (see Section 4 for details). See Appendix Table A9 for robustness to alternative control 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 9: Employment Effects of Successful Elections on Firms' Other Establishments

Panel A. All Establishments and Three-Digit NAICS Establishments

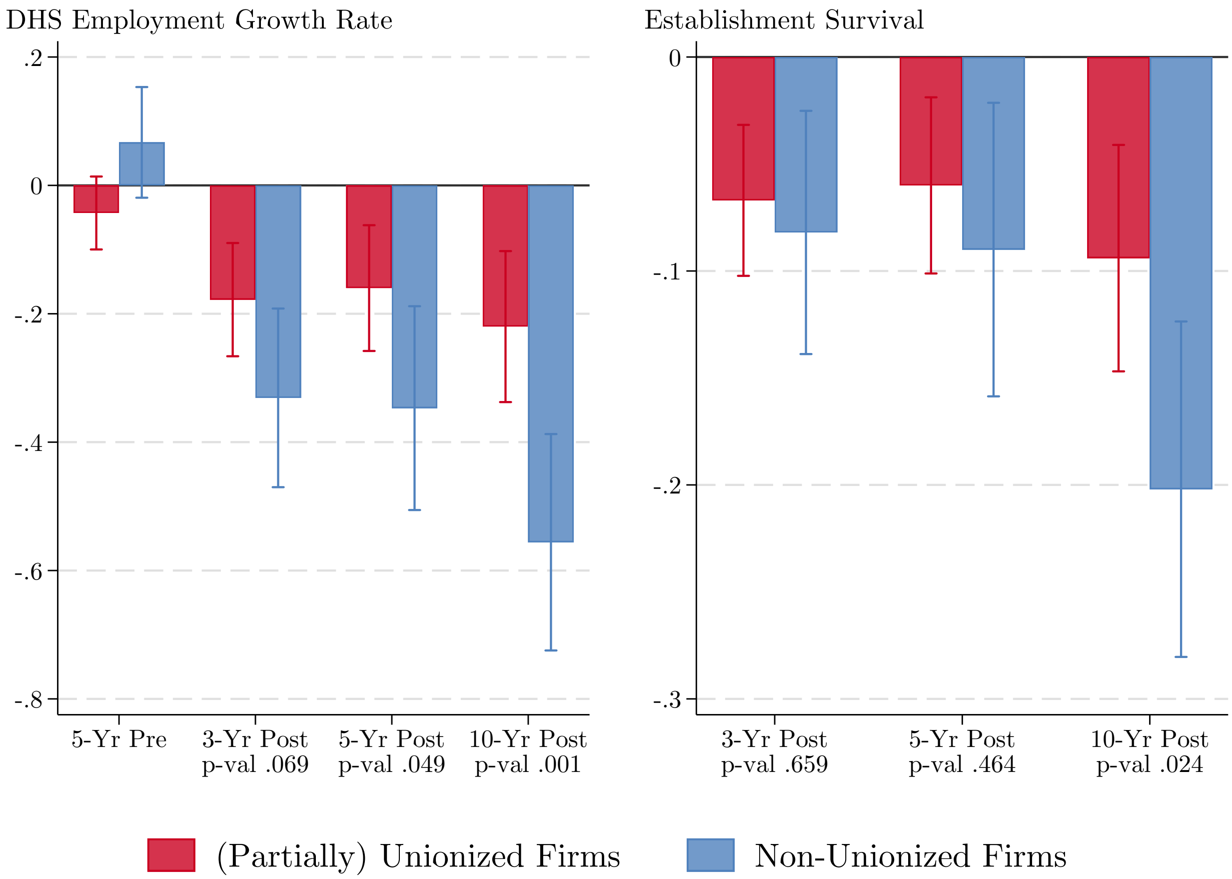


Panel B. Estimates by Elections' Employment Share



Note: This figure plots the δ_n coefficients from estimating specification 8. The sample is manufacturing establishments at multi-establishment firms where *another* establishment experienced a union election. See Appendix Section D.1 for details about the sample construction. The relative time and vote-share variables are defined based on the election at the firm's *other* establishment. We weight the regression by the observation's share of total firm-level employment across all establishments included in the sample in 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 establishments are part of MUs) or for establishments' previous contract status. Since we match establishments based on the election year, the industry variable is also from the year of the election. The "All Manufacturing Plants" estimates in the top panel include all manufacturing establishments with at least two employees during the year of the election. The "Within-NAICS 3 Plants" estimates restrict the sample to establishments that are in the same 3-digit NAICS industry as the election establishment. The bottom panel includes just 3-digit NAICS industry matches but separately estimates the effects by whether the election establishment comprised more than ten percent of the firm's employment in the same three-digit NAICS industry during the year of the 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 that we interact with event time.

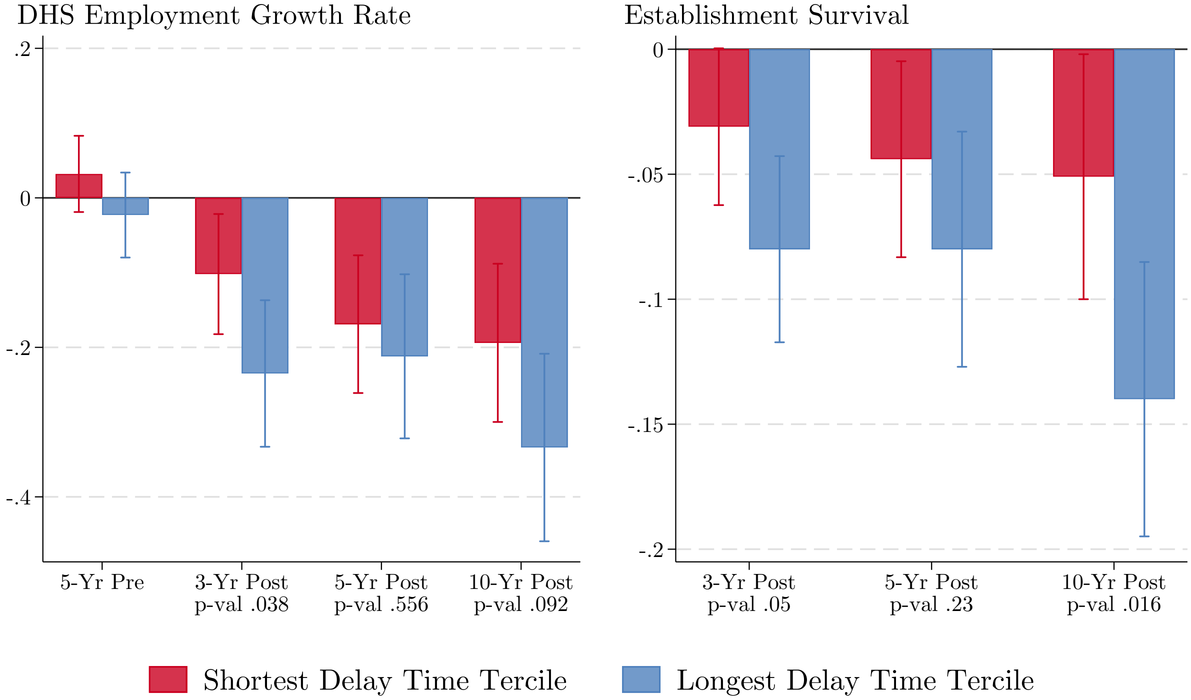
Figure 10: Unionized versus Non-Unionized Firm Heterogeneity



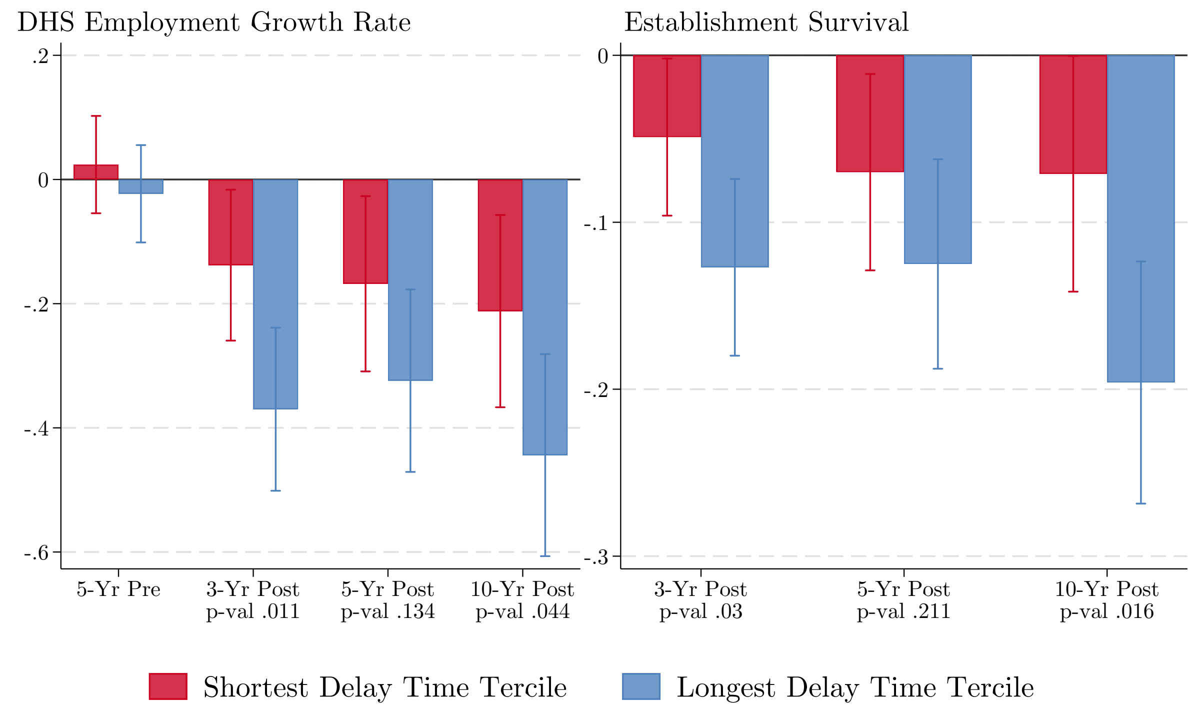
Note: This figure plots similar heterogeneity results as Figure 8 except that the heterogeneity is for elections at multi-establishment firms with at least one unionized establishment versus firms without any unionized establishments. See Appendix D for how we define firms' unionization status. Additionally, we directly include these heterogeneity groups interacted with cohort and event time as part of the controls.

Figure 11: Election Delay Heterogeneity

Panel A. All Manufacturing Elections



Panel B. Manufacturing Elections at Multi-Establishment Firms



Note: These figures plot the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 11 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 Appendix Section D 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 from -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 of 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 manufacturing 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.

10 Tables

Table 1: Winning versus Losing Election Establishment Summary Statistics

	All Industries		Manufacturing	
	Union Loses	Union Wins	Union Loses	Union Wins
Establishment Characteristics				
Employees	154	137	167	148
Payroll/Worker ($ 2019)	49,400	49,700	49,900	48,700
Establishment Age	9.65	10.0	9.82	9.58
Multi-Establishment Firm	0.512	0.476	0.525	0.464
Previous Contract at Establishment	0.090	0.147	0.088	0.152
Survival Base Rates				
5-Year Survival	0.818	0.765	0.847	0.779
10-Year Survival	0.667	0.610	0.702	0.608
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: Employment and Survival Estimates by Industry, 20–80% Vote-Share Elections

Industry Group:	Manufacturing		Services		Other Blue-Collar and Industrial	
Outcome:	DHS Emp	Survival	DHS Emp	Survival	DHS Emp	Survival
5-Year Pre Election	0.005		0.010		0.011	
	(0.015)		(0.012)		(0.016)	
2-Year Pre Election	-0.013		0.017*		-0.009	
	(0.012)		(0.009)		(0.012)	
5-Year Post Election	-0.174***	-0.047***	-0.057**	-0.026***	-0.192***	-0.058***
	(0.029)	(0.012)	(0.024)	(0.010)	(0.030)	(0.013)
10-Year Post Election	-0.231***	-0.075***	-0.059**	-0.017	-0.229***	-0.083***
	(0.033)	(0.015)	(0.027)	(0.012)	(0.033)	(0.015)
Industry + Employment Ctrls.	X	X	X	X	X	X
Flexible Ctrls.	X	X	X	X	X	X
Industry Group Number of Elections	6000	6000	8000	8000	5000	5000
Industry Group Share of Elections	0.302	0.302	0.414	0.414	0.284	0.284

Note: This table presents the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 11 for elections in three different broad industry groups. Manufacturing is defined as NAICS sectors 31–33, services is defined as NAICS 44–45 and 51–81, and other is defined as 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.

Table 3: Election Delay Heterogeneity, Continuous Delay Time Specification

Treatment:	Log Delay Time				Residualized Log Delay	
Outcome:	DHS Emp	Survival	DHS Emp	Survival	DHS Emp	Survival
3-Year Post Election	-0.124**	-0.057**	-0.117**	-0.046**	-0.120**	-0.055**
	(0.058)	(0.023)	(0.055)	(0.021)	(0.059)	(0.023)
5-Year Post Election	-0.121*	-0.064**	-0.103*	-0.052**	-0.113*	-0.060**
	(0.063)	(0.026)	(0.060)	(0.025)	(0.065)	(0.026)
10-Year Post Election	-0.147**	-0.071**	-0.152**	-0.073**	-0.132*	-0.061*
	(0.073)	(0.033)	(0.072)	(0.032)	(0.074)	(0.033)
Industry + Employment Ctrls.	X	X	X	X	X	X
Flexible Ctrls.	X	X			X	X
Number of Elections	6,000	6,000	6,000	6,000	6,000	6,000

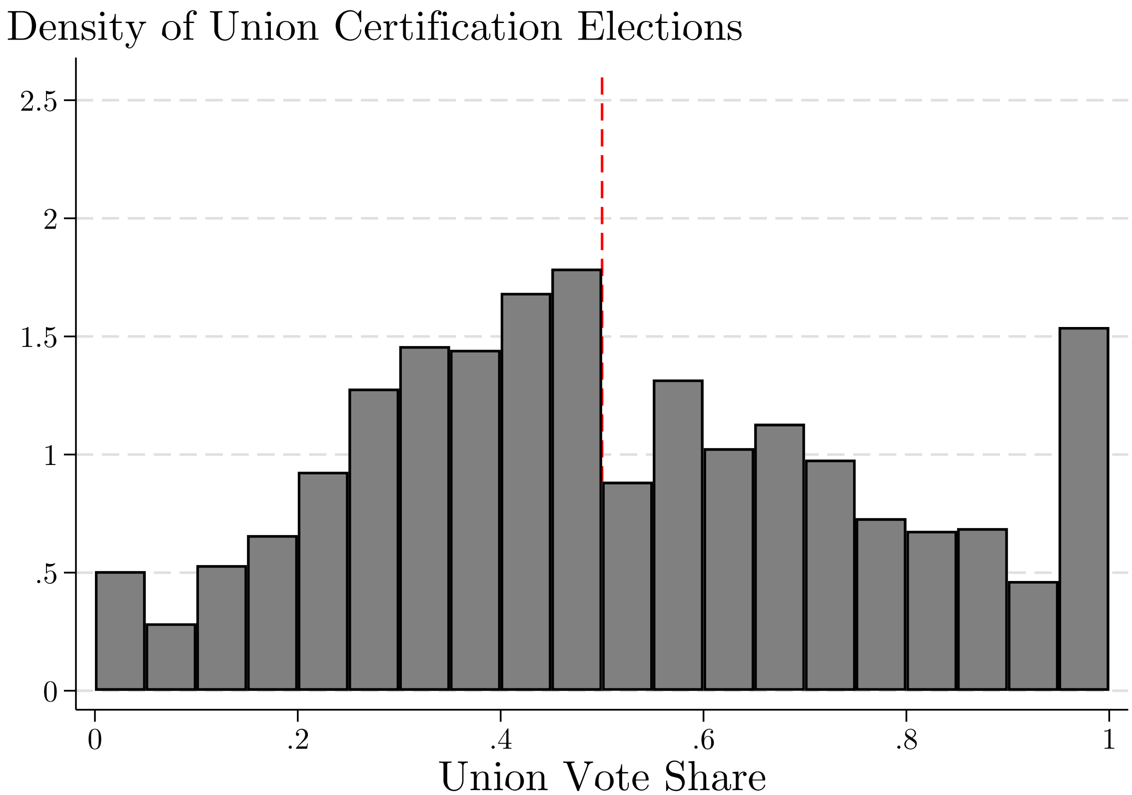
Note: This table presents coefficient estimates from a modified version of specification 8. Specifically, we interact the treatment by event time indicators with the continuous *log delay time*. See Appendix Section D for details on how we calculate the delay time. The table reports the coefficients on these interactions at various time horizons. Thus, a survival coefficient of -0.05 means that the effect of successful unionization on survival is 0.5 pct. pts. lower for elections with a ten percent 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's 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).

Appendix For Online Publication:

Unionization, Employer Opposition, and Establishment Closure

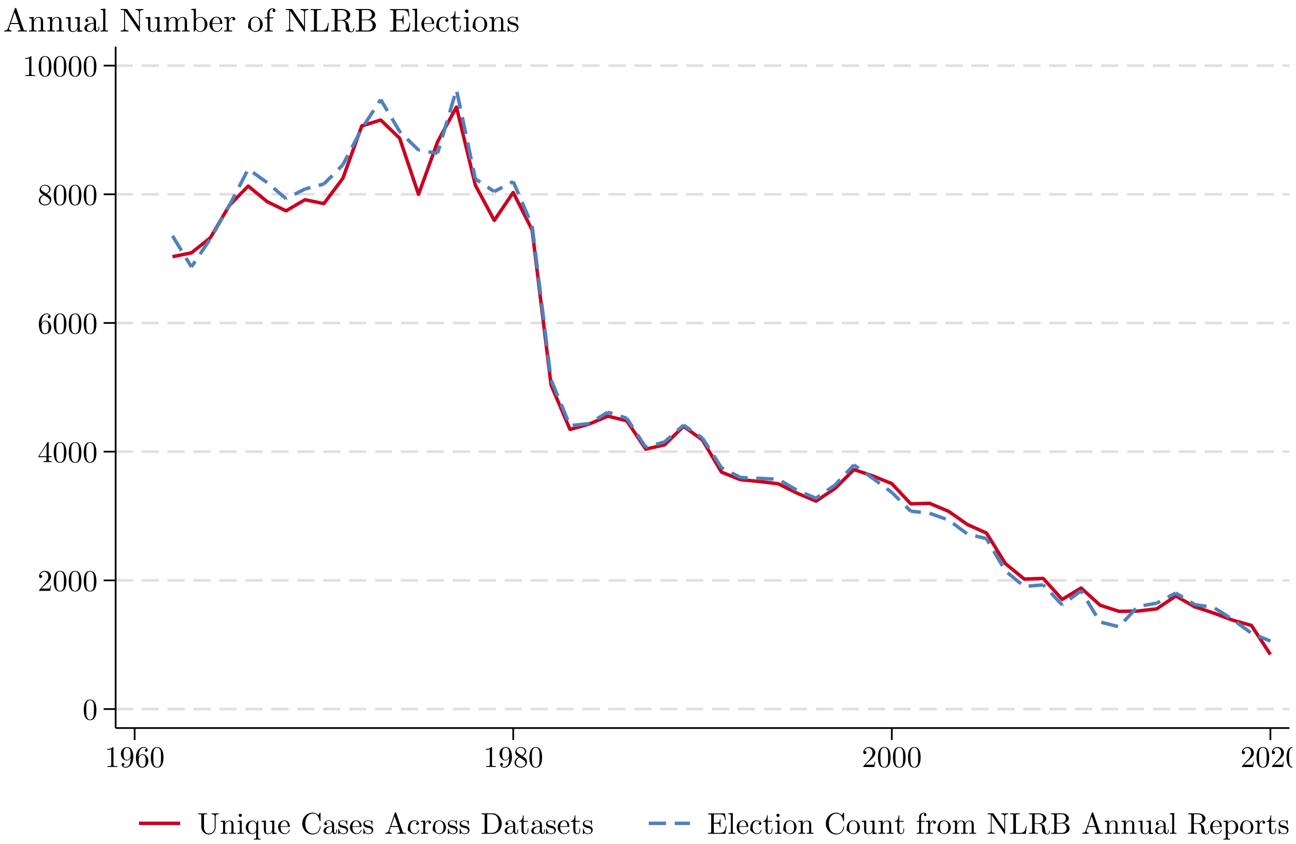
A Appendix Figures

Figure A1: Election Vote-Share Histogram, All In-Sample Elections



Note: This figure plots the vote-share histogram of elections included in our sample. 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 D for details). Panel A of Figure 1 plots the vote-share distribution for elections with 50 + votes to better illustrate the vote-share manipulation in this setting. See Frandsen (2017) for evidence of manipulation using formal tests that accommodate discrete running variables.

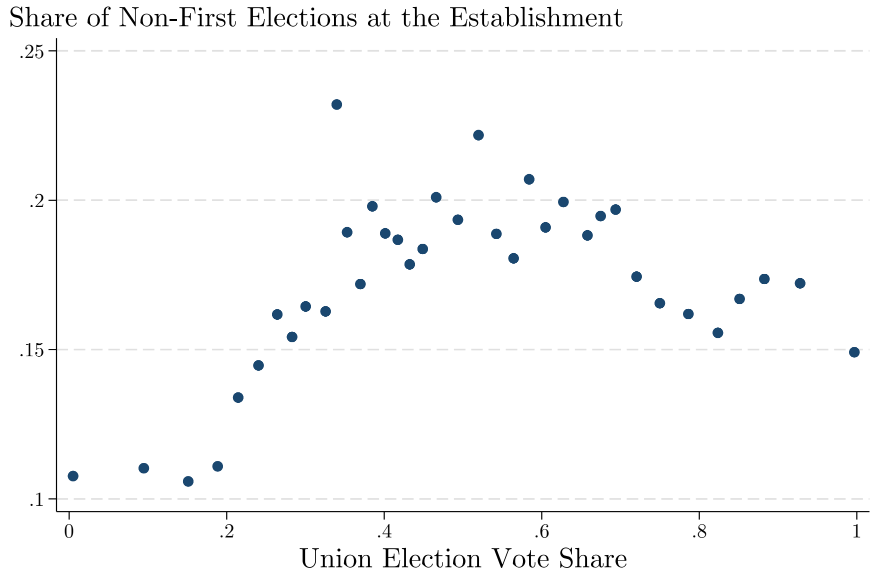
Figure A2: Number of Unique Case Numbers Across Datasets versus NLRB Annual Reports



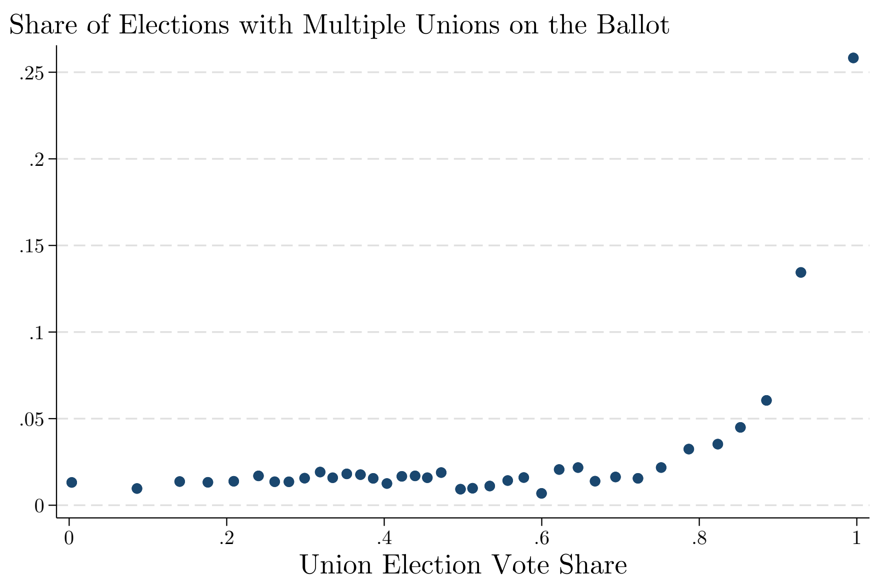
Note: This figure plots the total number of unique NLRB election cases each year in our dataset and in the annual NLRB reports. These include all case types (e.g., RC cases and non-RC cases). We create our dataset by combining union election datasets from Henry Farber, J.P. Ferguson, and Thomas Holmes and publicly available data from the NLRB and picking one observation for each NLRB case number. See Appendix D for details on our data construction process.

Figure A3: Characteristics of Extreme Low- and High-Vote-Share Elections

Panel A. Share of Non-First Elections at the Establishment by Vote Share



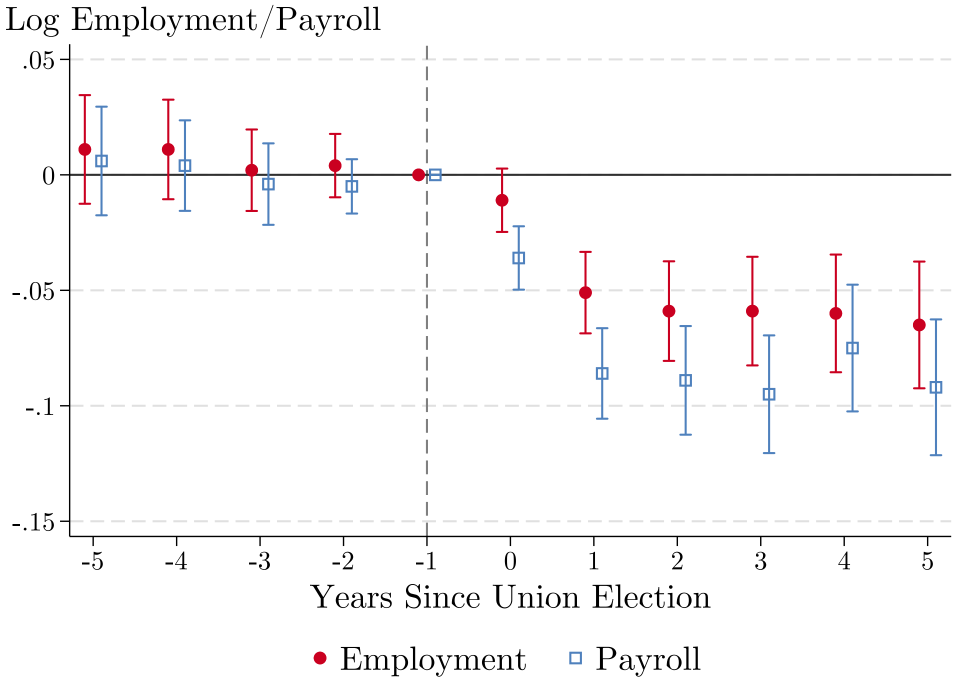
Panel B. Share of Multi-Union Elections by Vote Share



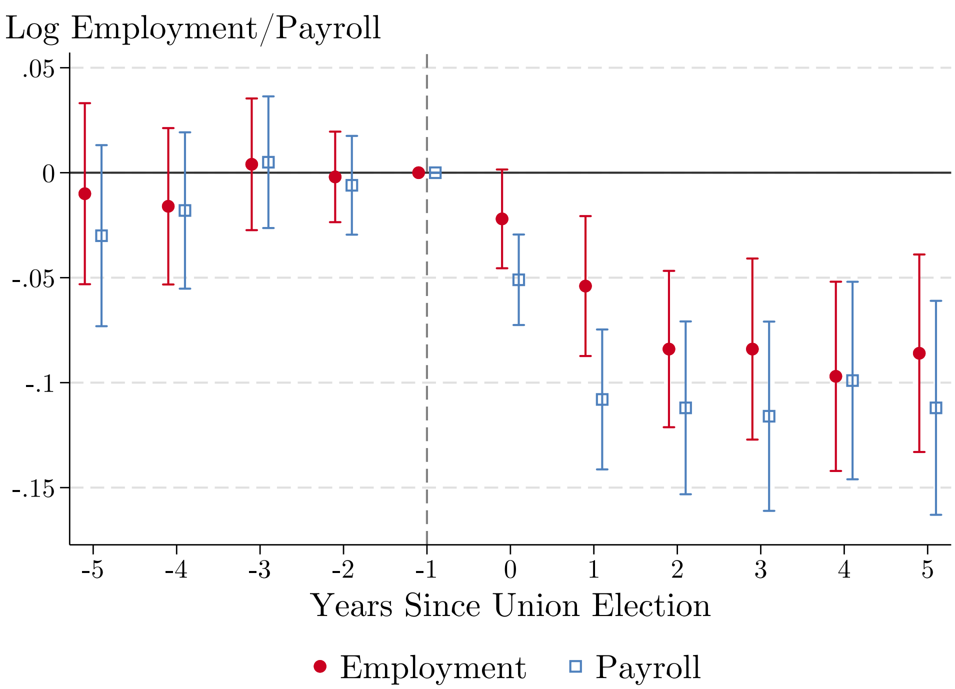
Note: These two figures plot how election characteristics differ across the vote-share distribution. They highlight that very low vote-share (e.g., around 0–20%) and very high vote-share (e.g., around 80–100%) elections often differ based on ex-ante characteristics from the other elections. Panel A. plots the share of elections that are not the first election at the establishment (e.g., a second election following an initial loss). To define "non-first" elections, we use name, city, and state matching with the external Census data rather than using the exact LBD matches with our Census data sample. Panel B. plots the share of elections with multiple unions on the ballot. The sample for both panels is based on our "external elections dataset" described in Appendix D. For these figures, the sample construction is slightly different. Specifically, we generally exclude non-first elections and elections with multiple unions on the ballot from the sample. Since these are the outcomes of interest, we include these elections in the sample. However, for the non-first-election figure, we do exclude multi-union elections and for the multi-union election figure, we exclude non-first elections.

Figure A4: Log Employment and Payroll Estimates, 20–80% Vote-Share Elections

Panel A. All Union Elections

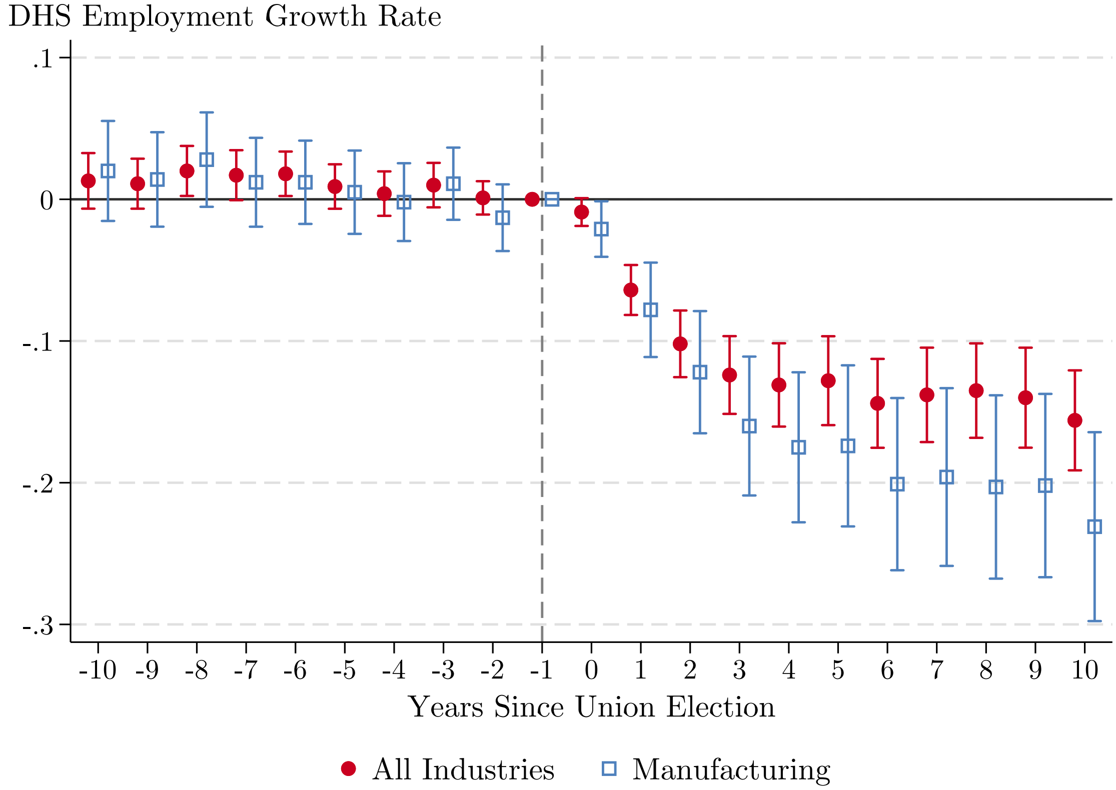


Panel B. Manufacturing Union Elections



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.

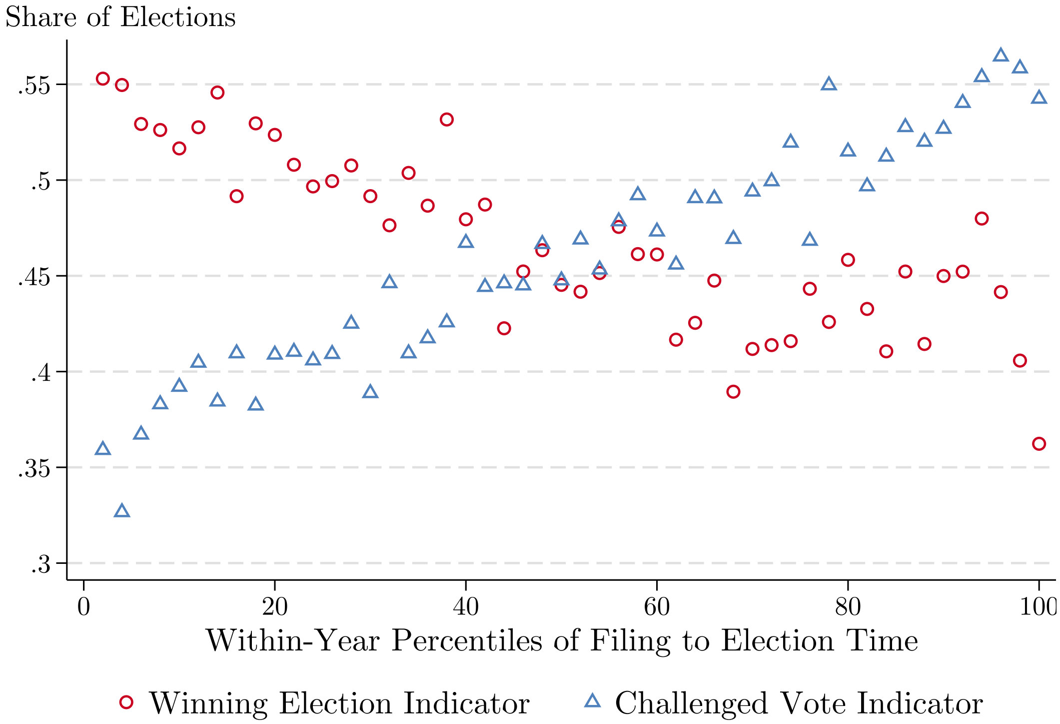
Figure A5: DHS Employment Estimates, 20–80% Vote-Share Elections, 10 Yr Pre- and Post-Periods



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. However, 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 estimate includes fewer elections than the -5 estimate). Consequently, each of the -5 to -10 point estimates averages over slightly different cohorts. This is why we use the -5 to 10-year estimates for the main analysis.

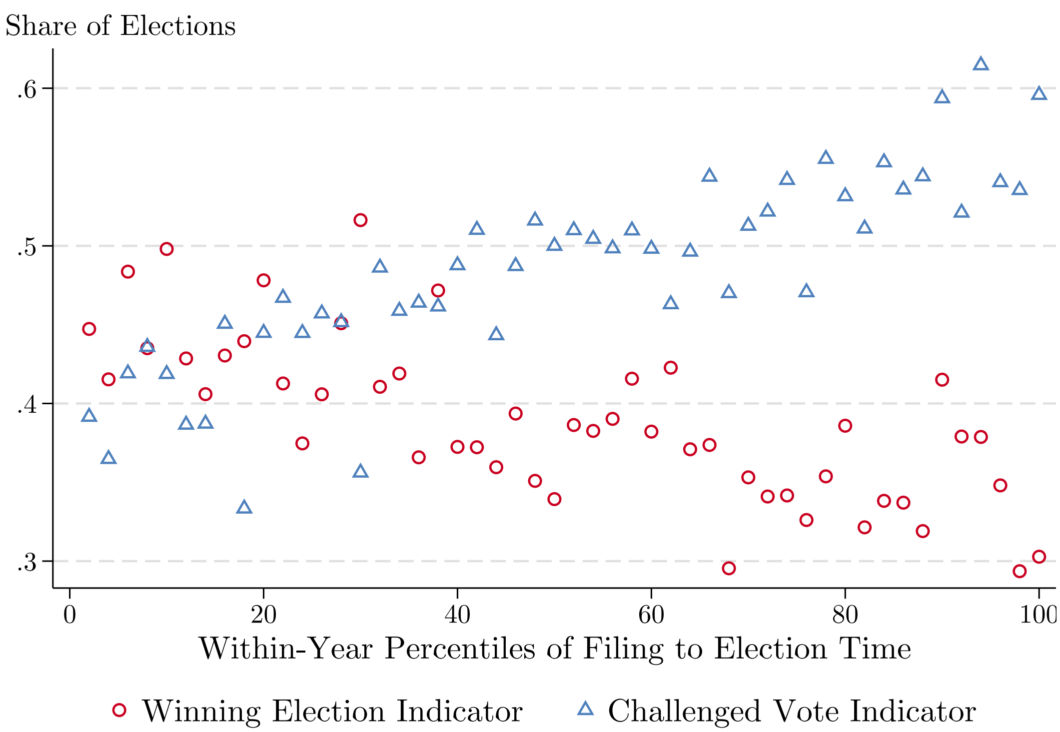
Figure A6: Election Win Rates and Challenged Vote Rates by Delay Time

Panel A. All Elections



Conditional Regression Coefficients x 100: Winning = -.081 (.007). Challenged = .067 (.007).

Panel B. Manufacturing

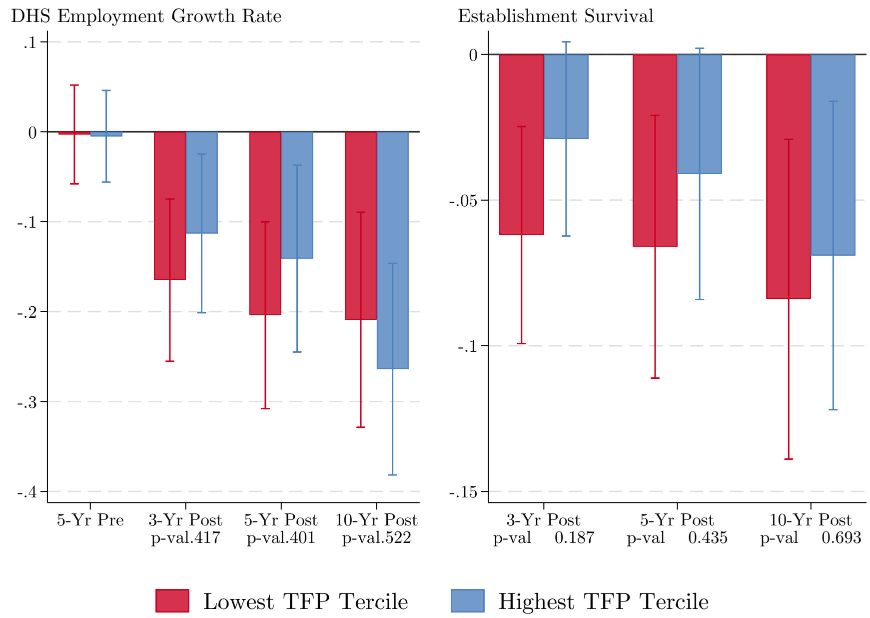


Conditional Regression Coefficients x 100: Winning = -.07 (.014). Challenged = .058 (.014).

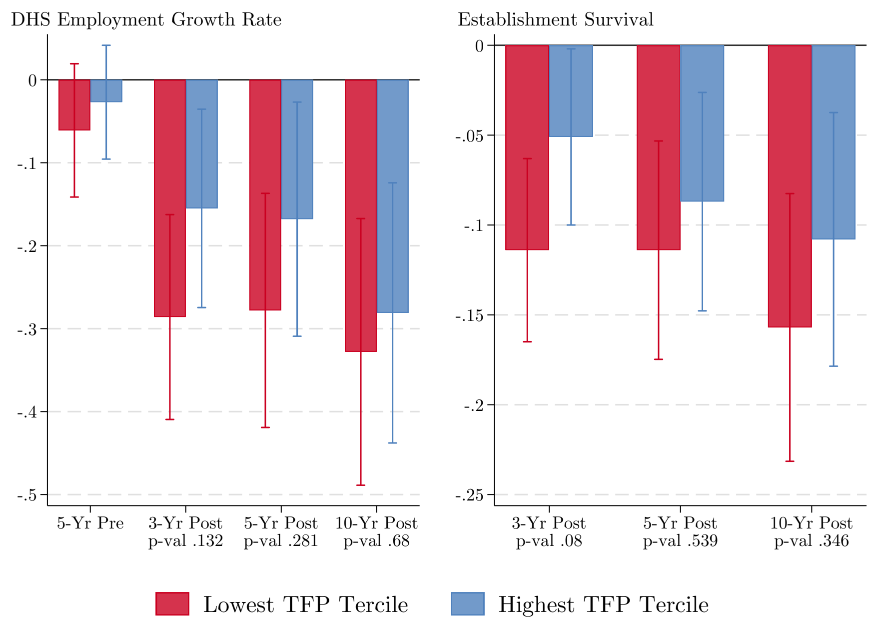
Note: This figure plots the relationship between pre-election delay times, election win rates, and challenged votes in elections. Pre-election delay times are defined as the number of days between the election petition being filed and the election date. We then take the within-year percentiles of the 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 D. 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.

Figure A7: Establishment-Level Total Factor Productivity Heterogeneity

Panel A. All Manufacturing Elections



Panel B. Elections at Multi-Establishment Firms



Note: These figures plot the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 11 for elections in different terciles of the baseline TFP distribution. These terciles are defined based on establishments' pre-election cost-share-based productivity measures from the Annual Survey of Manufactures (ASM) calculated by Cunningham et al. (2022). The TFP terciles are defined based on within-year and within six-digit NAICS productivity rankings. See Appendix D for details. We plot the coefficients for the first and third terciles, but we estimate effects for all three terciles and a fourth group of establishments where TFP is missing. The sample includes all manufacturing union elections with 20–80% vote shares inclusive. It includes observations from -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). Panel A reports the results for all manufacturing elections. Panel B further restricts the sample to just elections at multi-establishment firms.

B Appendix Tables

Table A1: Union Election Matched Sample Construction

	All Elections		Winning Elections	
	Elections	Eligible Voters	Elections	Eligible Voters
Panel A: NLRB Election Sample				
All Election, 1981-2005	94,824	5,991,865	44,492	2,461,138
Representation Elections (RC)	77,349	5,111,675	39,397	2,071,859
$>$ 5 Eligible Voters	69,789	5,084,061	34,247	2,053,210
Non-Contested Elections	66,353	4,590,121	31,378	1,668,877
Panel B: Final NLRB Sample Industry Shares				
Manufacturing	0.307	0.408	0.253	0.307
Other	0.266	0.186	0.263	0.177
Services	0.426	0.405	0.484	0.515
Panel C: Matched Census Sample				
Elections Matched to Census Establishments	46,000			
Final Establishment-Level Outcome Sample	27,000			
20-80% Election Sample	19,000			

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 with five or fewer 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. 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 2. The shares plotted in each row are the share of total elections or eligible voters from each of the specific broad industry groups. 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 $t = -1$.

Table A2: Narrow-Bandwidth and Within-Firm Identification Strategy Results

Industry Group:	40-60% Vote Share Bandwidth		Within-Firm Identification	
Outcome:	DHS Emp	Survival	DHS Emp	Survival
5-Year Pre Election	-0.012	–	-0.033	–
	(0.021)	–	(0.051)	–
2-Year Pre Election	-0.003	–	0.004	–
	(0.010)	–	(0.022)	–
3-Year Post Election	-0.096***	-0.038***	-0.195***	-0.048***
	(0.023)	(0.009)	(0.049)	(0.018)
5-Year Post Election	-0.088***	-0.037***	-0.207***	-0.066***
	(0.026)	(0.011)	(0.052)	(0.023)
Year X Cohort FEs	X	X		
Firm X Year X Cohort FEs			X	X
Number of Elections	6,000	6,000	1,000	1,000

Note: This table presents the δ_n coefficients (i.e., the interaction between winning a union election and being n years from the election) from estimating specification 8 for two alternative samples and sets of controls. The outcome variable for the first and fourth columns is establishment-level DHS employment growth relative to time -1. The outcome variable for the second and fifth columns is establishment-level log employment. The outcome variable for the third and sixth columns is an indicator for whether the establishment exists at time t. The "40-60% Elections" columns are from estimating specification 8 only including elections in all industries with vote shares between 40–47.5% and 52.5-60%. For this specification, we only include year-by-cohort fixed effects as controls. The "Within-Firm" columns are from estimating specification 8 only including elections at firms that had at least one winning and losing election in the same year, after imposing all of the other sample restrictions for our main strategy. For this specification, we only include firm-by-year-by-cohort fixed effects. 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 A3: Manufacturing versus Services Employment and Survival Estimates, Robustness Checks

Specification:	Baseline		Ind + Emp Ctrls		Good Matches		> 25% Barg Unit Share		30-70%	
Outcome:	DHS Emp	Survival	DHS Emp	Survival	DHS Emp	Survival	DHS Emp	Survival	DHS Emp	Survival
5-Year Difference	-0.117***	-0.021	-0.117***	-0.020	-0.144***	-0.024	-0.121***	-0.028*	-0.132***	-0.026
	(0.037)	(0.016)	(0.035)	(0.015)	(0.044)	(0.018)	(0.041)	(0.017)	(0.043)	(0.018)
10-Year Difference	-0.172***	-0.058***	-0.164***	-0.051***	-0.196***	-0.061***	-0.193***	-0.06***	-0.171***	-0.054**
	(0.043)	(0.019)	(0.040)	(0.018)	(0.05)	(0.022)	(0.047)	(0.021)	(0.049)	(0.022)
Industry + Employment Ctrls.	X	X	X	X	X	X	X	X	X	X
Flexible Ctrls.	X	X			X	X	X	X	X	X

Note: This table presents robustness results for the differences between the service-sector and manufacturing results in Table 2. 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 2. The "Ind + Emp Ctrls." columns only include baseline industry and employment controls as described in Section 4. The "Good Matches" columns restrict the sample to election matches to which we give at least a 95% rating (see Appendix D for details). The "Barg Unit Share" columns restrict the sample to elections where the bargaining unit is at least 25% of the total establishment employment. The 30–70% columns restrict the sample 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 A4: Borusyak et al. (2022) Imputation Estimator Results

Estimator:	Stacked TWFE				Borusyak et al. Imputation Estimator			
Sample:	All Industries		Manufacturing		All Industries		Manufacturing	
Outcome:	DHS Emp	Log Emp	DHS Emp	Log Emp	DHS Emp	Log Emp	DHS Emp	Log Emp
5-Years Pre Election	0.018	0.015	0.012	-0.010	0.018	-0.007	0.012	-0.034
	(0.012)	(0.011)	(0.022)	(0.021)	(0.012)	(0.011)	(0.022)	(0.021)
2-Years Pre Election	-0.000	0.003	-0.016	-0.005	-0.000	0.002	-0.016	-0.007
	(0.006)	(0.006)	(0.011)	(0.011)	(0.006)	(0.006)	(0.012)	(0.011)
3-Years Post Election	-0.131***	-0.070***	-0.170***	-0.099***	-0.122***	-0.051***	-0.169***	-0.071***
	(0.013)	(0.011)	(0.024)	(0.020)	(0.013)	(0.013)	(0.025)	(0.022)
5-Years Post Election	-0.135***	-0.075***	-0.188***	-0.108***	-0.131***	-0.047***	-0.193***	-0.072***
	(0.015)	(0.013)	(0.027)	(0.022)	(0.015)	(0.015)	(0.028)	(0.025)
Industry + Employment Ctrls.	X	X	X	X	X	X	X	X

Note: This table compares estimates from implementing our main "stacked TWFE" specification and the imputation-style estimator from Borusyak et al. (2022). The first four columns present coefficient estimates that are identical to the estimates presented in Figures 3 and 4. The treatment effect estimates in the last four columns are for the same outcomes and controls as the first four columns but estimated using the `did_imputation` Stata package. We include a five-year pre-period, so the estimates are based on assuming that parallel trends hold for the five years leading up to the union election. The pre-trend estimates in the first columns follow Borusyak et al. (2022)'s recommendation to estimate unit, time, and covariate coefficients using only untreated observations in a *separate* specification and including a certain number of leads of treatment. However, we deviate from their recommendation by omitting the event time $n = -1$ coefficient, rather than the earliest group of time periods. We make this change because (1) DHS employment growth rates are mechanically missing for $n = 1$, and (2) it makes the pre-trend estimates more comparable to the results from our main strategy. Note that the point estimates for DHS employment growth pre-trends are identical between our stacked TWFE and the imputation-style estimator. Since the DHS employment growth rates are calculated using long differences relative to event-time $n - 1$ and these specifications do not include establishment-level fixed effects, the pre-trend estimates are already equivalent to only estimating the pre-trend coefficients using untreated observations.

Table A5: Employment and Survival Bargaining Unit Share Interaction, 20–80% Vote-Share Elections

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

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 while accounting for overall post-election trends across all elections by bargaining unit share. A survival estimate of -0.05 means that increasing the share of the establishment covered by the bargaining unit by 10% leads to an additional 0.5 pct. pct. increase in establishment exit.

Table A6: Pre-Election Employment Growth Trends by Vote Share, 20–80% Elections

Outcome:	DHS Employment Growth Rate			
Industry Group:	All Industries		Manufacturing	
5-Year Pre Election × Vote Share	0.050	0.033	0.029	-0.018
	(0.037)	(0.025)	(0.069)	(0.047)
4-Year Pre Election × Vote Share	0.018	0.019	0.026	0.008
	(0.032)	(0.024)	(0.059)	(0.045)
3-Year Pre Election × Vote Share	0.028	0.029	0.022	0.018
	(0.023)	(0.023)	(0.041)	(0.042)
2-Year Pre Election × Vote Share	0.006	0.012	-0.026	-0.018
	(0.018)	(0.019)	(0.035)	(0.037)
Industry + Employment Ctrls.	X	X	X	X
Flexible Ctrls.		X		X
Number of Elections	19,000	19,000	6,000	6,000

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 ρ coefficients from equation 12). A five-year coefficient of 0.03 implies that elections with 75% vote shares grew approximately 1.5% slower during the five years before the election than an election with 25% vote shares. 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).

Table A7: Post-Election Outcome Trends by Vote Share, 20–80% Vote-Share Elections

Outcome:	DHS Emp Growth Rate		Establishment Survival	
Panel A: All Industries				
Event-Time × 0-50% Vote Share – θ Estimates				
3-Year Post Election	-0.216**	-0.085	0.013	0.036
	(0.095)	(0.103)	(0.036)	(0.040)
5-Year Post Election	-0.220**	-0.066	-0.031	0.004
	(0.110)	(0.122)	(0.045)	(0.049)
10-Year Post Election	-0.332***	-0.193	-0.080	-0.038
	(0.125)	(0.140)	(0.054)	(0.060)
Event-Time × 50-100% Vote Share – $\theta + \tau$ Estimates				
3-Year Post Election	-0.280**	-0.286**	-0.028	-0.029
	(0.131)	(0.131)	(0.052)	(0.052)
5-Year Post Election	-0.381**	-0.389***	-0.052	-0.053
	(0.149)	(0.149)	(0.063)	(0.063)
10-Year Post Election	-0.271*	-0.278*	-0.071	-0.073
	(0.164)	(0.164)	(0.073)	(0.073)
Panel B: Manufacturing				
Event-Time × 0-50% Vote Share – θ Estimates				
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)
Event-Time × 50-100% Vote Share – $\theta + \tau$ Estimates				
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 + Employment Ctrls.	X	X	X	X
Flexible Ctrls.	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 12 and capture linear trends in post-election outcomes for losing elections. The `Event-Time` × `50-100` rows present estimates of $\theta + \tau$ 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 A8 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.

Table A8: Post-Election Outcome Trends by Vote Share, 20–80% Vote-Share Elections, Employment and Industry Ctrls.

Industry Group:	All Industries		Manufacturing	
Outcome:	DHS Emp	Survival	DHS Emp	Survival
3-Year Post Election				
Event-Time × 0-50% Vote Share (θ)	-0.134	0.021	-0.181	0.042
	(0.100)	(0.037)	(0.159)	(0.059)
Event-Time × 50-100% Vote Share ($\theta + \tau$)	-0.361***	-0.052	-0.543**	-0.035
	(0.126)	(0.051)	(0.250)	(0.099)
5-Year Post Election				
Event-Time × 0-50 % Vote Share (θ)	-0.119	-0.009	-0.150	0.023
	(0.116)	(0.047)	(0.187)	(0.076)
Event-Time × 50-100% Vote Share ($\theta + \tau$)	-0.450***	-0.085	-0.537*	-0.033
	(0.141)	(0.060)	(0.275)	(0.116)
10-Year Post Election				
Event-Time × 0-50% Vote Share (θ)	-0.218	-0.052	-0.186	-0.020
	(0.133)	(0.057)	(0.224)	(0.097)
Event-Time × 50-100% Vote Share ($\theta + \tau$)	-0.354**	-0.107	-0.676**	-0.209
	(0.157)	(0.070)	(0.309)	(0.140)
Exclude 50% Elections	X	X	X	X
Industry + Employment Ctrls.	X	X	X	X
Flexible Ctrls.				
Number of Elections	19,000	19,000	6,000	6,000

Note: This table presents the same estimates as in Table A7 but only includes the baseline industry and employment controls.

Table A9: Single- Versus Multi-Establishment Firm Heterogeneity, Robustness Checks

Specification:	Baseline		Ind + Emp Ctrls		30-70%	
Outcome:	DHS Emp	Survival	DHS Emp	Survival	DHS Emp	Survival
5-Year Difference	-0.068	-0.061**	-0.061	-0.062***	-0.034	-0.057**
	(0.058)	(0.024)	(0.047)	(0.020)	(0.065)	(0.027)
10-Year Difference	-0.149**	-0.093***	-0.119**	-0.091***	-0.067	-0.06*
	(0.066)	(0.03)	(0.052)	(0.024)	(0.075)	(0.034)
Industry + Employment Ctrls.	X	X	X	X	X	X
Flexible Ctrls.	X	X			X	X

Note: This table presents robustness results for the differences between single- and multi-establishment firms presented in Figure 8. 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 8. The "Ind + Emp Ctrls." columns only include baseline industry and employment controls 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 A10: Effects of Successful Elections on Firms' Other Establishments

Outcome:	DHS Employment	Survival
1-Year Post Election	0.017	0.006
	(0.016)	(0.006)
2-Year Post Election	0.043**	0.012
	(0.019)	(0.008)
3-Year Post Election	0.044**	0.022**
	(0.021)	(0.009)
4-Year Post Election	0.048*	0.015
	(0.025)	(0.010)
5-Year Post Election	0.034	0.023**
	(0.025)	(0.011)
Industry X Year X Cohort FEs	X	X
Baseline X Cohort Controls	X	X

Note: This table presents the DHS employment growth rate and survival estimates that are estimated as described for Figure 9. The DHS employment growth rate estimates exactly match the DHS employment estimates presented in that table.

Table A11: Unionized versus Non-Unionized Firm Heterogeneity, Robustness Checks

Specification:	Baseline		Ind + Emp Ctrls		Contracts since 1990		30-70% Elections	
Outcome:	DHS Emp	Survival	DHS Emp	Survival	DHS Emp	Survival	DHS Emp	Survival
5-Year Difference	-0.187**	-0.03	-0.142	-0.021	-0.179	-0.057	-0.197*	-0.029
	(0.095)	(0.041)	(0.089)	(0.038)	(0.112)	(0.049)	(0.108)	(0.046)
10-Year Difference	-0.336***	-0.108**	-0.278***	-0.090**	-0.412***	-0.149**	-0.305**	-0.098*
	(0.104)	(0.048)	(0.097)	(0.045)	(0.121)	(0.058)	(0.119)	(0.055)
Industry + Employment Ctrls.	X	X	X	X	X	X	X	X
Flexible Ctrls.	X	X			X	X	X	X

Note: This table presents robustness results for the differences between multi-establishment firms with and without any unionized establishments presented in Figure 10. 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 10. The "Ind + Emp Ctrls." columns only include baseline industry and employment controls as described in Section 4. The "Contracts since 1990" column only classifies firms as unionized versus non-unionized starting in 1990. This provides us with at least five years of pre-election FMCS contract data for all firms, which we can use to determine 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 NLRB Elections, Wages, and Productivity Literature Review

One motivation for the analysis in Section 6 is that the large survival and employment effects we document seem at odds with the existing evidence of muted effects of successful union elections on wages and productivity. In this section, we review the prior literature on both outcomes and discuss the degree to which the evidence supports wage or productivity effects as potential causes of the survival effects we document.

Recent Unionization and Wage Increases The most relevant analyses of the effects of recent union elections on wages are a series of regression discontinuity papers that find *small* effects on workers' wages. Frandsen (2021) implements a regression discontinuity analysis that accounts for the non-random selection just around the 50% threshold. He estimates the effect of unionization in all industries on worker-level quarterly earnings changes one year following the election and finds no increase in workers' earnings. Sojourner et al. (2015) also analyze the wage effects of nursing-home unionization using worker-level data, but their smaller sample size yields imprecise overall wage estimates. With this in mind, their decile-specific analysis finds that unionization has a large *negative* earnings impact on workers with the highest pre-election earnings and zero effects across the rest of the distribution. Similarly, DiNardo and Lee (2004) and LaLonde et al. (1996) do not find any impact of unionization on average payroll per worker at the establishment level.[58] Finally, Freeman and Kleiner (1990b) compare the wages at establishments with successful union elections to "their closest competitors," identified by the firms themselves. They find that successful elections lead to, at most, small wage increases but that successful elections do lead to large changes in personnel practices (e.g., grievance procedures and seniority provisions). The one exception to this literature that has struggled to find an effect of unionization on workers' compensation is Knepper (2020). Using regression discontinuity and difference-in-differences approaches, Knepper (2020) mirrors the previous research by finding that successful elections do not lead to increases in workers' average wages. However, he also finds that a successful election at one establishment leads to very large increases in non-wage benefits across the *entire firm*.

Overall, the above literature is inconsistent with the idea that newly certified unions drive firms out of business by raising wages. However, there are a few caveats to this interpretation. First, most of the above papers only look at the relatively short-run impact of unionization (e.g., up to one year following the election). Consequently, it is possible that longer-run wage increases drive the survival effects we estimate. However, such longer-run effects are inconsistent with the fact that our employment and payroll estimates are relatively similar even five years following the election. This implies that we do not estimate long-run increases in average payroll per worker (although these estimates do not account for changes in worker composition). Second, the above regression discontinuity papers only analyze the wage effects of very close union elections and may not extrapolate to elections that win by larger margins of victory. However, we find large survival

[58] Since DiNardo and Lee (2004)'s survival and employment results differ from ours and Frandsen (2021)'s, we interpret their other estimates with some caution.

and employment effects even for very close union elections where these regression discontinuity papers do not find wage increases. Overall, the difficulty this literature has had finding positive wage effects from recent union elections motivates our exploration of non-wage reasons why unionization decreases establishment survival. However, the literature cannot completely rule out the survival effects being driven by long-run wage increases or costly increases in non-wage benefits.

Another body of literature that may seem at odds with the idea that recent unionization has not led to wage increases is the "union wage premium" literature. This literature estimates the wage premium that unionized workers receive relative to non-union workers using cross-sectional or panel data on workers. These papers generally find a union wage premium of 10–20% (Lewis, 1986; Card, 1996; Farber et al., 2021).

However, there are several ways to reconcile "the union wage premium" with the smaller establishment-level estimates of successful union elections. First, the quasi-experimental literature on the effects of union elections only considers recent elections since the 1980s. The union wage premium, however, also includes workers at establishments that were unionized before then. Given the drastic changes in the state of labor relations (Kochan et al., 1986a) and the macroeconomic environment (Bluestone and Harrison, 1982) during the 1980s, it is plausible that unionization before the 1980s led to large wage increases but unionization afterward had a smaller impact on wages. Supporting this, Freeman and Kleiner (1990b) argue that one reason they did not find that unionization in the 1980s led to increased wages was "the unfavorable economic environment of the period: the decline in union representation, deregulation of industries, increased foreign competition, and high unemployment that are likely to have raised the elasticity of labor demand facing newly organized labor and the reduced the ability of the unions to raise wages." This story is also consistent with the union wage premium decreasing from around 20% in the 1980s to 10% in the 2010s (Farber et al., 2021).[59] Second, the union wage premium includes businesses unionized without an NLRB election (e.g., card check or voluntary recognition). If there are different causal effects of NLRB versus non-NLRB unionization or the selection into elections differs between these organizing modes, this could reconcile the union wage premium with the establishment-level estimates.

Finally, the union wage premium may be biased by two different selection issues that are often addressed in the analyses of union elections. The first bias is that there may be non-random selection of which workers become union members. For example, more productive workers may become union members or, as argued by Frandsen (2021), unions may lower the returns to skills and consequently attract lower-skilled workers. The second bias is that there may be non-random selection into which establishments are unionized (e.g., more productive establishments that would pay high wages anyway are more likely to unionize).[60]

[59] Charles et al. (2023) provide evidence that the rise of Chinese imports (i.e., the "China Shock") contributed to some of these negative trends for unions during this time period.

[60] Several papers in the union premium literature address the worker selection issue and argue that the union wage premium is not driven by this selection (see e.g., Lemieux (1998) and Krashinsky (2004) although de Chaisemartin and D'Haultfoeuille (2020) finds that worker-level selection may be severe). The union wage premium literature, however, generally does not address the establishment-level selection into unionization. Dinlersoz et al. (2017)'s finding that more productive establishments attract union elections suggests that this selection may be severe and the causal effect of unions on wages may be overstated by the union wage premium.

Unionization and Productivity The most relevant studies about unionization and productivity are a series of recent quasi-experimental studies analyzing the impact of unionization on establishment-level productivity. These papers generally find that unionization has a null or positive impact on productivity. For example, Sojourner et al. (2015) implement a regression discontinuity analysis and find that unionization at nursing homes decreases employment with no impact on the quality of care, which they interpret as productivity increases. Hart and Sojourner (2015) and Dube et al. (2016) analyze recent elections using difference-in-differences designs and find that unionization does not decrease student achievement at charter schools and that unionization improves patient outcomes at hospitals, respectively. Similarly, DiNardo and Lee (2004) find no impact on output per worker for elections in manufacturing (although the caveat in Footnote 58 applies here too) and LaLonde et al. (1996) find that unionization in manufacturing has no effect on output per total hours although it decreases output per employee.

Recent non-U.S. quasi-experimental evidence mirrors the previous findings by showing a positive impact of unions on productivity (Barth et al., 2020). Finally, several older papers use cross-sectional comparisons to compare the productivity of more versus less unionized locations or industries. While these estimates are mixed, reviews of this literature conclude that it generally finds small positive or zero effects (Freeman and Medoff, 1984; Kuhn, 1998; Hirsch, 2004). For example, Kuhn (1998) writes that "Most [productivity] estimates are positive, with the negative effects largely confined to industries and periods known for their conflictual union-management relations, or to the public sector."

Overall, numerous null and positive productivity estimates previously discussed suggest that productivity is unlikely to be driving the large establishment survival effects that we document. Again, however, a few caveats prevent us from completely ruling out the productivity explanation. First, like the wage literature, most estimates are relatively short-run and may not capture longer-run productivity decreases. Second, some of the recent quasi-experimental work on the impact of unionization on productivity focuses on industries where we do not find significant negative survival effects (e.g., nursing homes, education, and hospitals). This may be because the survival effects we document in some industries raise problems with studying the effect on productivity in these industries by comparing the surviving establishments (Lee, 2009). Finally, much of the recent literature does not analyze the effect on total factor productivity (TFP) but instead looks at the effects on various proxies for productivity like output per worker or product quality (see Brown and Medoff (1978) for a discussion of analyzing the effect of unionization on TFP versus other productivity proxies).

There are also several related literatures that may seem to imply that recent unionization decreases productivity, but such conclusions require additional assumptions. First, Holmes (1998) finds "an abrupt increase in manufacturing activity when one crosses a state border from a" right-to-work state to a non-right-to-work state. However, in this case, right-to-work laws represent a bundle of "pro-business" policies, so the results do not imply that unionization by itself reduces manufacturing employment. Second, Krueger and Mas (2004) and Mas (2008) show that strikes at Bridgestone/Firestone and Caterpillar led to large productivity decreases.[61] However, strikes, especially of that size and duration, have become increasingly uncommon since 1984, which suggests that the productivity declines from potential strikes are unlikely to explain the exit effects we document.[62] Finally, Galdon-Sanchez and Schmitz (2002) and Schmitz (2005) document how unionized firms increase their productivity in response to increases in competition. However, this evidence does not provide direct evidence that unions decrease productivity, but instead shows that some unionized firms can increase productivity by changing work practices.

D Data and Matching Details Appendix

NLRB Union Election Data

Union Election Data Sources 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. For more details about the sources of these data, see JP Ferguson's website here.

NLRB Election Case Numbers The ID variable in the election data is an `NLRB Case ID Number.` This case number is assigned after an election petition is first filed. A single case number, however, could include multiple different vote counts. For example, there might be (1) multiple different tallies of the same election or (2) multiple elections for the same case number.[63] Additionally, there might be separate elections for multiple different bargaining units filed under the same case number (e.g., if a union initially filed a petition for one bargaining unit but the NLRB then split the bargaining unit). Consequently, it is important to pick the vote count that actually corresponds to the outcome of the certification election. Finally, since the different data sources cover overlapping time periods, we have multiple observations of the same case number in different datasets.

We deal with multiple observations per case number within datasets differently for the different data sources. For the public NLRB data (the "Public Data"), there is information indicating why there are multiple observations for a single case number. Consequently, for a given bargaining unit, we pick the final tally of the last election for each case number. This ensures that we take the vote tally that determines the unions' certification for cases where there are multiple counts of the same election or multiple ordered elections for the same bargaining unit. 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

[61]See also, Gruber and Kleiner (2012) on the negative effects of hospital strikes on patient health.

[62] According to the FMCS work stoppage data, the Bridgestone and Caterpillar strikes studied by (Krueger and Mas, 2004) and Mas (2008) lasted 314 and 531 days, respectively. Compared to the other 13,905 strikes in the FMCS data since 1984, these were the 5th and 27th largest strikes considering the number of workers on strike and the strike duration. Alternatively, for the strikes following successful union elections in our sample, the median (mean) duration of work stoppages was only 28 (70) days. Additionally, Krueger and Mas (2004) find the largest decreases in product quality right before the strike when contentious bargaining was occurring, and when the replacement and striking workers were working side-by-side. This evidence is more consistent with general adversarial labor relations leading to productivity declines rather than the direct costs of strikes.

[63]There could be multiple tallies for the same election due to challenged votes (e.g., the first tally would not include challenged votes while the final tally would include challenged votes that were determined to be valid). There could be multiple elections for the same case number if an NLRB director orders a second election due to objections to the first election

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 with one observation per case number within each dataset but duplicates across datasets. We take one observation per case number across datasets. For picking a single case number per dataset, we deprioritize observations in the Farber data due to irregularities in those data. Additionally, we prioritize the public data because we have more confidence that we are picking the correct observation across duplicates within the same case number.

Variables in the Election Dataset We define the following variables from the union election data that we use for our analysis and for our matching algorithm

- **Election City, State, and Address:** The data contain the city and state of the election that we use to match each election to an establishment in the LBD. For many observations, we also observe a street address that we also use for the matching.

  For the "public data", we observe an address for the employer and an address for the election site. There are two conceptual reasons why these addresses might be different. First, the election might not be held at the employers' location.[64] This suggests that the employer's address is better for name and address matching to Census establishments. Second, the listed address for the employer might be a corporate headquarters rather than the establishment where the bargaining unit works. This suggests that the election address is better for name and address matching. Since it is not conceptually clear which address to use, we check which address is more likely to match the text in the bargaining unit description (e.g., "all warehousemen at its Louisville, KY facility"). We find that the election site address is more likely to match the address information in the bargaining unit description. Consequently, we use the election site addresses when the two address fields disagree.

- **Election Vote Shares:** We define election vote shares as the number of votes for the union divided by the total number of votes in the election. This differs from the adjusted vote shares constructed in DiNardo and Lee (2004) and Frandsen (2021) to address the "integer problem" with constructing vote shares[65]. We do not apply this adjustment for two reasons. First, the integer problem is especially problematic for regression discontinuity designs but less of an issue with our difference-in-differences design. Second, since we don't impose any restrictions on the number of votes cast in the election, the adjustment proposed in DiNardo and Lee (2004) would lead to larger changes in our vote shares (e.g., a six-person election would be adjusted from 50% to 41.7%).

[64]For example, when strikes, pickets, or lockouts are in progress, the election may be held at a neutral location (NLRB, 2020). As another example, when the employer's location is different than the employees' worksite (e.g., security guards), the election might be held at the worksite

[65]The integer problem refers to the fact that since vote shares are based on a discrete number of votes, there will be a mechanical discontinuity in the number of elections with exactly 50% vote shares

- **Contested Elections:** We define contested elections as elections with multiple unions on the ballot. We drop these elections for two reasons. First, these elections are often "union raids" where one union already represents a specific bargaining unit and another union challenges that union for representation (Sandver and Ready, 1998). Consequently, a winning election, in this case, would not lead to a switch from the establishment being non-unionized to unionized but instead just a switch in which union represents the bargaining unit. Second, the reported vote totals for multi-union elections may not actually represent the workers' support for the union. In particular, for multi-union elections, if none of the options (e.g., "union 1", "union 2", or "no union") receive the majority of the votes, a runoff election is held between the highest two options (Fraundorf, 1990). Consequently, the unions' true support (the union vote share from the first election) may be different than the unions' support in the observed runoff election results.

- **Election Industry:** The election data contain industry codes indicating the industry of the election analysis. For our main analysis, we use the Census industry codes for the establishments we match each election to. For some of our analysis of the unmatched NLRB data (e.g., Figures 1 and A6 and Tables A1), we use the election industry codes to split up manufacturing and non-manufacturing elections. Since the industry codes in the election data come from different vintages (e.g., SIC versus NAICS industry codes), we use the modal employment-weighted industry crosswalks from Eckert et al. (2020) to crosswalk the industry codes to consistent NAICS 2012 industry codes.

- **Bargaining Unit Size and Share of Total Employment:** We define the `bargaining unit size` as the number of eligible voters from the NLRB election data. We define the `bargaining unit share of total employment` as the bargaining unit size divided by the establishment-level employment one year for the union election. Since we do not impose that the bargaining unit is smaller than the establishment, we cap the share at one.

- **Election Filing Date:** We define treatment timing based on the date that the election was filed. To maximize the number of observations that we observe election filing dates for, we pull the dates across case numbers when some observations are missing from one dataset (e.g., if the filing date is only available for a case in the Ferguson data but not the Farber data, we pull date from the Ferguson to Farber data). For five percent 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. However, as described above, we pull both dates across datasets when they are missing for some observations. 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).

FMCS Contract Data

We combine contract data from Thomas Holmes for 1984–2003 and from the FMCS for 1997–2019. The Homes data are available here and the FMCS data are partially available here and the rest were obtained via a FOIA request. They include both notices of *initial contracts* (i.e., first-contract negotiation after an election) and *contract renegotiation or reopening* for existing contracts. There are two reasons that these contract notices likely underrepresent the universe of unionized establishments in the U.S. First, 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. Second, some contract notices may represent a contract covering multiple establishments but we only match each contract to one establishment.[66]

There are duplicate observations both across the Holmes and FMCS datasets and within each dataset.[67] However, unlike the NLRB election data, we have no IDs to restrict the dataset to unique observations. Consequently, to deal with duplicates, we match all contract observations to the Census establishments in the LBD and drop duplicates when multiple contract observations match to the same Census establishment.

We use the contract data to define

- **Previous contract at an establishment:** for each election establishment, we define an indicator for whether the establishment has a previous FMCS contract ever matched to the same establishment (e.g., indicating that another bargaining unit was already unionized at this establishment). To avoid contract matches related to the union election, we only include matched contracts starting one year before the election.

- **Unionized versus Non-Unionized 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. For the unionized versus non-unionized firm heterogeneity check, we also include elections at establishments with a previous contract (defined above) as unionized firms.

[66]Sometimes, the FMCS contract notices explicitly mention that they apply to multiple locations (e.g., the address indicating various locations). In these cases, we will still only match the contract notice to one establishment if there is alternative location data available.

[67]The across-dataset duplicates come from the fact that the datasets overlap. The within-dataset duplicates could come from an employer and union submitting an FMCS notice for the same contract.

FMCS Work Stoppage Data

For the analysis in Footnote 62, we use work stoppage data from the FMCS from 1984–2005. The data are available here. They include both strikes and employer-initiated lockouts. We match the work stoppages to the election data based on exact company names and cities rather than the *Soft TF-IDF* algorithm we use for the main analysis. Prior to matching, we use the same cleaning algorithms described below to clean the employer and city names in the FMCS work stoppage data.

Longitudinal Business Database

In Section 3, we mention potential concerns with how the LBD allocates employment across establishments at multi-establishment firms that could bias our results. To be more precise about the issue, while the LBD is an establishment-level dataset, some of the employment and payroll input data are received at higher levels of aggregation (e.g., at the EIN level). For example, one source used to construct the LBD is *IRS form 941* which provides 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. This model primarily imputes employment changes across establishments based on their past employment. Consequently, employment changes at an establishment that is part of a multi-establishment firm might initially be allocated across all establishments. Thus, the LBD would initially underestimate the establishment-level decrease in employment. To correct some of these mistakes, the Census receives establishment-level information from the Company Organization Survey (COS), Economic Censuses, and Annual Survey of Manufactures (ASM) that provide more accurate measures of establishment-level employment and survival. These alternative surveys are not, however, conducted for all establishments annually (e.g., the Economic Census is only conducted every five years). So there might be a few years lag before the LBD reports the correct establishment employment and exit. This lag mirrors the spike in establishment births and deaths every five years during the economic census years when the Census has establishment-level data for each establishment (Jarmin and Miranda, 2002). See Chow et al. (2021) for details about these issues with the LBD construction.

We use the LBD to define the following establishment-level variables

- **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).

- **Establishment-Level NAICS Codes:** We classify each establishment into a 2012 NAICS industry using the Fort and Klimek (2016) NAICS codes.

Establishment-Level TFP from the Annual Survey of Manufactures

We define establishment-level productivity using inputs and outputs from the Annual Survey of Manufactures (ASM) and TFP measures calculated by Cunningham et al. (2022) following the work of Foster et al. (2016). To classify each election into different terciles of the establishment 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 establishment'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 it's $E_i - 2$ 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 Establishments

Our data on union elections and contract notices contain information on the name and location of the employer, but no unique identifiers (like EIN) that could be used to directly link the establishments to administrative Census observations. We instead use a fuzzy-matching algorithm to link each election or contract to its corresponding Census record from the Standard Statistical Establishment List/Business Register. The algorithm is based on the name and geographic similarity of establishments. Our algorithm is based upon the Soft TF-IDF approach used by Kline et al. (2019), but extends their approach to incorporate the additional address data.

Name and Address String Cleaning: We start by standardizing and cleaning the name and address strings. Our cleaning procedure builds on the `stnd_compname` and `stnd_address` Stata name standardization programs (Wasi and Flaaen, 2015). We clean addresses as follows:

1. Remove most symbols, non-numeric or letter characters, and non-standard ASCII characters.
2. Removed PO boxes, building/suite/room numbers, and company names at the start of addresses (e.g., **GENERAL SUPPLY COMPANY** 2651 1ST STREET.)
3. Standardize common address and city name strings (e.g., `ST` ⇒ `STREET`, `TWENTY FIRST` ⇒ `21ST`, and `LIC` ⇒ `LONG ISLAND CITY`) and correct common address and city misspellings.

We clean the employer names as follows

1. Remove most symbols, non-numeric or letter characters, and non-standard ASCII characters.
2. Remove the portion of company names in parentheses. The union election data often contain supplemental information in the parentheses portion of the name (e.g., `(wage employees only)`).

3. Remove the portion of company names following `DOING BUSINESS AS (DBA)` or `A DIVISION OF`

4. Combine consecutive singleton letters and symbols separated by spaces (e.g., `A T & T` ⇒ `AT&T` and `D R HORTON` ⇒ `DR HORTON`).

5. Remove company entity types (e.g., `CORP`, `INC`, etc.), articles, and standard common company names (e.g., `MANUFACTURERS` ⇒ `MANUFACTURING`).

Election, Contract, and Census Address Geocodes: We geocode all the addresses. This allows us to construct measures of address similarity based on the geographic distance between two addresses. We use geographic distance rather than string distance to measure address similarity because there may be addresses with very similar strings that are very different addresses (e.g., `100 Main St.` may be very far away from `10 Main St.`).

For the election and contract data, we first try to geocode all addresses with the Census Bureau's Geocoding API because these geocodes are the most likely to match the Census's internal geocodes. For the observations where the Census's geocoder cannot find a geocode, we try the geocodio geocoder. When a street address is missing or cannot be geocoded, we use the geocode for the city/state or the zip code.

For the Census data, we use the geocodes in the SSEL/Business Register (DeSalvo et al., 2016). These geocodes, however, are only available since 2002 (Akee et al., 2017). For observations where we do not have a geocode, we first try to match it to a geocoded address. If the same address was not geocoded from 2002–2016, we instead take the average geocode of all addresses we see in 2002–2016 in the same city/state or zip code.

Matching Algorithm We implement a matching algorithm based on the string similarity of the cleaned employer names and the geographic distance between geocoded addresses. The standard Soft TF-IDF algorithm computes a match score between two firm names that is increasing in their string similarity. The algorithm is particularly suitable for our application since it overweights similarities in uncommon words between the two names and discounts similarities in common words. 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"). Consequently, we also incorporate the geography information to distinguish between these potential matches.

We implement our matching algorithm as follows

1. For each election, we take all Census establishments in the same state that share at least one common word.[68]

[68] We require that the establishments share at least one common word because this vastly reduces the number of string and distance calculations we need to make. For single-word companies, we only require that the potential matches share the same first letter. This allows us to match single-word establishments even with misspellings

2. For each election-establishment pair, we calculate the Soft TF-IDF similarity measure between the employer name strings. Specifically, let A_j be the set of all words in the election name string and B_k be the set of all words in the establishment name string. The total number of election names is J, and the total number of Census names is K. The Soft TF-IDF distance is defined as

$$s_{jk} = \text{Soft TF-IDF}(A_j, B_k) = \sum_{w \in A_j} \text{weight}(w, A_j) \times \text{m-score}(w, B_k) \tag{A1}$$

where weight(w, A_j) is defined as

$$\text{weight}(w, A_j) = \text{TF}(w, A_j) \times \text{IDF}(w, A, B) \Bigg/ \left[\sum_{w' \in A_j} \left(\text{TF}(w, A_j) \times \text{IDF}(w, A, B)\right)^2 \right]^{1/2} \quad \text{where} \tag{A2}$$

$$\text{TF}(w, A_j) = \frac{\text{freq}(w, A_j)}{\sum_{w' \in A_j} \text{freq}(w', A_j)} \quad \text{and} \tag{A3}$$

$$\text{IDF}(w, A, B) = -1 \times \log \left(\frac{\sum_{j'} \mathbb{1}[w \in A_{j'}] + \sum_{k'} \mathbb{1}[w \in B_{k'}]}{J + K} \right). \tag{A4}$$

Intuitively, the `TF` portion of the weight gives higher weights to words that are part of shorter names. The `IDF` portion of the weight gives higher weights to less common words relative to all words included in any election or Census establishment name. We give higher weights to less common words because two names sharing a common word (e.g., `manufacturing`) is less likely to indicate a correct match than two words sharing a less common word (e.g., `wanaque`).

The m-score(w, B_k) is defined as follows

$$\text{m-score}(w, B_k) = \overline{m}(w, B_k) \times \text{weight}(\overline{w}, B_k) \times \mathbb{1}[\overline{m}(w, B_k) > \theta] \tag{A5}$$

where $\overline{m}(w, B_k)$ is the highest Jaro-Winkler distance between the word w and any word in the name B_k

$$\overline{m}(w, B_k) = \max_{w' \in B_k} \text{Jaro-Winkler}(w, w') \tag{A6}$$

and $\overline{w}$ is the word in B_k that maximizes the Jaro-Winkler string distance. θ is a threshold below which the m-score is defined as zero. The Jaro-Winkler string distance is a measure of how similar two strings are. It considers the number of matching characters in the strings and the number of transpositions necessary to get the strings to match (e.g., `Boston` and `Bostno` require one transposition). Finally, it also places a higher weight on matching characters at the beginning of strings. See Kline et al. (2019) for details.

3. We calculate the Haversine distance between the election and Census establishment geocoordi-

nates as follows

$$d_{j,k} = \min(\text{Haversine Distance}(\text{geo\_coord}_j, \text{geo\_coord}_k), \overline{d}). \tag{A7}$$

where $\overline{d}$ is our distance top code (e.g., distances above a certain threshold are unlikely to be informative).

4. We combine the string similarity measure and the distance measure for each pair of elections and establishments as follows

$$\text{match score}_{jk} = (1 - \beta) \cdot s_{jk} + \beta \left[1 - (d_{jk}/\overline{d})^{\gamma}\right] \tag{A8}$$

where β is the relative weight placed on distance versus string name similarity. γ is the relative weight placed on very close versus farther away matches (e.g., a very concave γ places much more weight on exact geographic matches than matches that are even slightly farther away).

5. For each election, we pick the Census establishment with the highest match score$_{jk}$. This yields a potential match for each election but these matches may be very low quality or incorrect.

6. We only keep matches where match score$_{jk}$ is above a minimum threshold p.

The matching algorithm has several tuning parameters that determine the relative weights placed on each component of the final match score. For the parameters used to calculate the Soft TF-IDF score and the final match score (e.g., θ, the p parameter in the JW string distance, and γ), we use details about our institutional setting to optimize these parameters in a principled manner. We first optimize the Soft TF-IDF parameters by matching each election record to, at most, one contract record. We then choose the parameters that maximize the discontinuity in the likelihood that an election record has a matching contract record across the 50% vote-share threshold.

To pick the minimum match score p, we exploit the fact that the size of the election bargaining unit in the election data and the number of employees at the Census establishment give us information about whether the match is correct. In particular, having a larger bargaining unit than the number of workers at the establishment indicates an incorrect match.[69] We first 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 bargaining unit size 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 $\underline{m}$. We assume that records where the name and geographic location match exactly are "true" matches, which correspondingly allows us to estimate that a pair

[69]There may be cases of larger bargaining unit sizes than establishment employment that actually are correct matches. For example, there may be data mistakes in the bargaining unit size, or the measures may cover different time periods.

of records with a match score of s is matched correctly with probability:

$$p(s) = \frac{m(s) - \underline{m}}{m(1) - \underline{m}}. \tag{A9}$$

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.

D.1 Employment Shifting Analysis Sample Construction

For the employment shifting analysis described in Section 6, we construct the sample and conduct the analysis as follows:

1. For each cohort, we start by taking all elections in our 20–80% manufacturing sample at multi-establishment firms. We then take all of the firms' other establishments that existed during the year of the election. We drop *other* establishments with one or fewer employees during the year of the election. Additionally, we drop *other* establishments that are in the same county as the election establishment. The reason for this restriction is when a firm has multiple establishments in the same county, our matching algorithm may match the election to the wrong establishment. This would lead to a mechanical negative spillover effect. We exclude establishments that ever experienced an election so our "spillover estimates" are not contaminated by direct effects. Yet, this conditioning could selectively bias our sample. The most plausible mechanism, however, biases us *against* finding positive spillovers. Specifically, assume that successful elections lead to *more* future elections at a firm. Since elections occur at relatively fast-growing establishments and the establishment needs to survive to hold a future election, we would drop faster-growing establishments at firms with successful elections. This would bias our spillover estimates downward.

2. For each of these *other* establishments, we assign treatment based on the election outcomes at the election establishment. For firms with multiple elections in the same year, we prioritize winning elections. For example, if a firm had a winning and losing election in the same year, we would label the spillover establishments as being at a firm with a winning election. This should downward bias our spillover estimates. For choosing between the other election characteristics at firms with multiple elections (e.g., the election establishment's industry), we prioritize the election with the largest bargaining unit.

3. From this total set of *other* establishments, we restrict to three subsets of establishments.

    (a) **Manufacturing Establishments:** these are the set of matching establishments that have a manufacturing industry code during the year of the election.

    (b) **Same Industry Establishments:** These are the set of matching establishments that are in the same three-digit NAICS industry as the election establishment during the year of the election.

(c) **High Employment Share Elections:** These are the elections that match to elections where the election establishment makes up at least 10% of the firm's total employment during the year of the election. Since we just use the employment share split for the same-industry sample, we define the election establishment's employment share based just on the total firm-level employment in the same three-digit NAICS industry.

4. For each matched *other* establishment, we calculate DHS employment growth rates relative to one year before the other establishment election. Note, we still require that the election establishment existed at least three years prior to the election, but do not require this restriction for the other establishments. However, to define the DHS employment growth rate, all establishments in the sample need to have non-zero employment one year before the election.

5. For all the cohorts in our sample, we stack together all of the "matched" establishments at the same firm during the year of the election. This construction results in some establishments being in the dataset 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. For this analysis, similar conditioning is more difficult because the Census `firm IDs` change over time, even for firms that stay in business, and establishments can switch to different `firm IDs`. This also motivates our two-way clustering by firm and establishment.

6. For this stacked sample of other establishments, we estimate a modified version of our main specification 8 with the following changes. We estimate this regression on the three specifications listed above. However, when we split the sample by the high-employment share elections, we pool the two groups together and pool the controls across the groups.

   (a) All relative time and vote-share variables are defined from the *other* establishment's election.

   (b) We weight the regression by each establishment's share of its firms' employment. The reason for the weighting is that the sample could include multiple establishments 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 establishments).

   (c) We two-way cluster the standard errors by the firm ID from the election year and establishment. Since establishments can change firm IDs and firms can experience multiple elections, these are not nested.

   (d) We include the same controls from our *flexible* controls specification except that the industry controls are defined based on the $t = 0$ rather than $t = -1$ values, and we do not include controls for previous contracts at the establishment.

7. For the regression weights, we only include employment at establishments in the sample in the denominator so that the weights sum to one. Specifically, the weight is the current

establishment's employment during the year of the election divided by the sum of employment during the year of the election across all other establishments that are in the regression sample. Consequently, when we restrict to only within-industry matches the denominator only includes employment at other establishments that are also in the same industry.

Appendix References

Akee, Randall, Elton Mykerezi, and Richard M Todd (2017) "Reservation Employer Establishments: Data from the U.S. Census Longitudinal Business Database," p. 35.

Barth, Erling, Alex Bryson, and Harald Dale-Olsen (2020) "Union Density Effects on Productivity and Wages," *The Economic Journal*, Vol. 130, pp. 1898–1936.

Bluestone, Barry and Bennett Harrison (1982) *The Deindustrialization of America: Plant Closings, Community Abandonment, and the Dismantling of Basic Industry*, New York: Basic Books, Inc., Publishers.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2022) "Revisiting Event Study Designs: Robust and Efficient Estimation," *Working Papers*, arXiv:2108.12419 [econ].

Brown, Charles and James Medoff (1978) "Trade Unions in the Production Process," *Journal of Political Economy*, Vol. 86, pp. 355–378.

Card, David (1996) "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis," *Econometrica*, Vol. 64, pp. 957–979.

de Chaisemartin, Clément and Xavier D'Haultfoeuille (2020) "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, Vol. 110, pp. 2964–2996.

Chow, Melissa, Teresa Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elisabeth Ruth Perlman, Martha Stinson, and T. Kirk White (2021) "Redesigning the Longitudinal Business Database," *NBER Working Paper*, p. w28839, Place: Cambridge, MA.

Cunningham, Cindy, Lucia Foster, Cheryl Grim, John Haltiwanger, Sabrina Wulff Pabilonia, Jay Stewart, and Zoltan Wolf (2022) "Dispersion in Dispersion: Measuring Establishment-Level Differences in Productivity," *Review of Income and Wealth*.

DeSalvo, Bethany, Frank Limehouse, and Shawn D. Klimek (2016) "Documenting the Business Register and Related Economic Business Data," *CES Working Papers 16-17*.

DiNardo, John and David S. Lee (2004) "Economic Impacts of New Unionization on Private Sector Employers: 1984-2001," *Quarterly Journal of Economics*, Vol. 119, pp. 1383–1441.

Dinlersoz, Emin, Jeremy Greenwood, and Henry Hyatt (2017) "What Businesses Attract Unions? Unionization over the Life Cycle of U.S. Establishments," *ILR Review*, Vol. 70, pp. 733–766.

Dube, Arindrajit, Ethan Kaplan, and Owen Thompson (2016) "Nurse Unions and Patient Outcomes," *ILR Review*, Vol. 69, pp. 803–833.

Eckert, Fabian, Teresa Fort, Peter Schott, and Natalie Yang (2020) "Imputing Missing Values in the US Census Bureau's County Business Patterns," *NBER Working Paper*, p. w26632, Place: Cambridge, MA.

Farber, Henry S, Daniel Herbst, Ilyana Kuziemko, and Suresh Naidu (2021) "Unions and Inequality Over the Twentieth Century: New Evidence from Survey Data," *Quarterly Journal of Economics*, Vol. 136, p. 169.

Fort, Teresa C and Shawn Klimek (2016) "The Effects Of Industry Classification Changes On Us Employment Composition," *CES Working Paper 18-28*, p. 24.

Foster, Lucia, Cheryl Grim, and John Haltiwanger (2016) "Reallocation in the Great Recession: Cleansing or Not?" *Journal of Labor Economics*, Vol. 34, pp. S293–S331.

Frandsen, Brigham (2021) "The Surprising Impacts of Unionization: Evidence from Matched Employer-Employee Data," *Journal of Labor Economics*, Vol. 39, p. 56.

Frandsen, Brigham R. (2017) "Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete," in Matias D. Cattaneo and Juan Carlos Escanciano eds. *Advances in Econometrics*, Vol. 38: Emerald Publishing Limited, pp. 281–315.

Fraundorf, Martha Norby (1990) "The effect of voting procedures on union representation elections: The multi-union case," *Journal of Labor Research*, Vol. 11, pp. 323–335.

Freeman, Richard B. and Morris M. Kleiner (1990) "The Impact of New Unionization on Wages and Working Conditions," *Journal of Labor Economics*, Vol. 8, pp. S8–S25.

Freeman, Richard and James Medoff (1984) *What Do Unions Do?*, N.Y.: Basic Books.

Galdon-Sanchez, José E. and James A. Schmitz (2002) "Competitive Pressure and Labor Productivity: World Iron-Ore Markets in the 1980's," *American Economic Review*, Vol. 92, pp. 1222–1235.

Gruber, Jonathan and Samuel A. Kleiner (2012) "Do Strikes Kill? Evidence from New York State," *American Economic Journal: Economic Policy*, Vol. 4, pp. 127–157.

Hart, Cassandra and Aaron J. Sojourner (2015) "Unionization and Productivity: Evidence from Charter Schools," *Industrial Relations*, Vol. 54, pp. 422–448.

Hirsch, Barry (2004) "What Do Unions Do for Economic Performance?" *Journal of Labor Research*, Vol. XXV, pp. 415–455.

Holmes, Thomas J. (1998) "The Effect of State Policies on the Location of Manufacturing: Evidence from State Borders," *Journal of Political Economy*, Vol. 106, pp. 667–705.

Jarmin, Ron S. and Javier Miranda (2002) "The Longitudinal Business Database," *SSRN Electronic Journal*.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar (2019) "Who profits from patents? rent-sharing at innovative firms," *Quarterly Journal of Economics*, Vol. 134, pp. 1343–1404.

Knepper, Matthew (2020) "From the Fringe to the Fore: Labor Unions and Employee Compensation," *The Review of Economics and Statistics*, Vol. 102, pp. 98–112.

Kochan, Thomas A, Katz Harry, and McKersie Robert (1986) *The Transformation of American Industrial Relations*, New York: Basic Books.

Krashinsky, Harry A. (2004) "Do Marital Status and Computer Usage Really Change the Wage Structure?" *The Journal of Human Resources*, Vol. 39, pp. 774–791, Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System].

Krueger, Alan and Alexandre Mas (2004) "Strikes, Scabs, and Tread Separations: Labor Strife and the Production of Defective Bridgestone/Firestone Tires," *The Journal of Political Economy*, Vol. 112.

Kuhn, Peter (1998) "Unions and the Economy: What we Know; What we Should Know," *The Canadian Journal of Economics*, Vol. 31, pp. 1033–1056.

LaLonde, Robert, Gerard Marschke, and Kenneth Troske (1996) "Using Longitudinal Data on Establishments to Analyze the Effects of Union Organizing Campaigns in the United States," *Annales d'Économie et de Statistique*, p. 155.

Lee, David S. (2009) "Training, Wages, And Sample Selection: Estimating Sharp Bounds On Treatment Effects," *The Review of Economic Studies*, Vol. 76, pp. 1071–1102.

Lemieux, Thomas (1998) "Estimating the Effects of Unions on Wage Inequality in a Panel Data Model with Comparative Advantage and Non-Random Selection," *Journal of Labor Economics*, Vol. 16, pp. 261–291.

Lewis, Gregg H. (1986) "Union Relative Wage Effects," in *Handbook of Labor Economics*, Vol. 2: Elsevier, pp. 1139–1181.

Mas, Alexandre (2008) "Labour Unrest and the Quality of Production: Evidence from the Construction Equipment Resale Market," *Review of Economic Studies*, Vol. 75, pp. 229–258.

NLRB (2020) "National Labor Relations Board: Casehandling Manual: Part Two Representation Proceedings."

Sandver, Marcus Hart and Kathryn J. Ready (1998) "Trends in and determinants of outcomes in multi-union certification elections," *Journal of Labor Research*, Vol. 19, pp. 165–172.

Schmitz, James A. (2005) "What Determines Productivity? Lessons from the Dramatic Recovery of the U.S. and Canadian Iron Ore Industries Following Their Early 1980s Crisis," *Journal of Political Economy*, Vol. 113, pp. 582–625, Publisher: The University of Chicago Press.

Sojourner, Aaron J., Brigham R. Frandsen, Robert J. Town, David C. Grabowski, and Min M. Chen (2015) "Impacts of Unionization on Quality and Productivity: Regression Discontinuity Evidence from Nursing Homes," *ILR Review*, Vol. 68, pp. 771–806.

Wasi, Nada and Aaron Flaaen (2015) "Record Linkage Using Stata: Preprocessing, Linking, and Reviewing Utilities," *The Stata Journal: Promoting communications on statistics and Stata*, Vol. 15, pp. 672–697.