Unionization, Employer Opposition, and Establishment Closure*

Sean Wang[†] Samuel Young[†] [‡]

December 12, 2022

Abstract

We study the effect of private-sector unionization on establishment employment and survival. Specifically, we analyze National Labor Relations Board union elections from 1981–2005 using administrative Census data. Our empirical strategy extends standard difference-in-differences techniques with regression discontinuity extrapolation methods. This allows us to avoid biases from only comparing close elections and to estimate treatment effects that include larger margin-of-victory elections. Using this strategy, we show that unionization decreases an establishment's employment and likelihood of survival, particularly in manufacturing and other blue-collar and industrial sectors. We hypothesize that two reasons for these effects are firms' ability to avoid working with new unions and employers' opposition to unions. We test this hypothesis for manufacturing elections and find that the negative effects are significantly larger for elections at multi-establishment firms. Additionally, after a successful union election at one establishment, employment increases at the firms' other establishments. Both pieces of evidence are consistent with firms avoiding new unions by shifting production from unionized establishments to other establishments. Finally, we find larger declines in employment and survival following elections where managers or owners were likely more opposed to the union. This evidence supports new reasons for the negative effects of unionization we document.

*We are grateful to Daron Acemoglu, David Autor, and Simon Jäger for guidance and advice throughout this project. We thank Josh Angrist, Jon Cohen, David Hughes, Sylvia Klosin, Tom Kochan, Felix Koenig, Mike Piore, Frank Schilbach, Garima Sharma, Martina Uccioli, John Van Reenen, Michael Wong, and Josef Zweimüller and participants at several seminars for helpful comments. This paper benefited greatly from Henry Hyatt and Kirk White's data expertise. We thank Stephanie Bailey, Jim Davis, and Nathan Ramsey for 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). All errors are our own. *Disclaimer: Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2389 (CBDRB-FY22-P2389-R9311 and CBDRB-FY22-P2389-R9358).*

Emails: sean.wang.econ@gmail.com and sammygyoung@gmail.com.

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

[‡]Arizona State University

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).[1] The standard economic explanation for why establishments close after unionization is that unions make establishments unprofitable by increasing wages or implementing other workplace changes. However, this explanation is not supported by existing research, which has found little evidence of successful union elections increasing wages or decreasing productivity.[2]

Consider two examples of how employers responded to unions that suggest alternative reasons why unionization may cause establishment closures. First, 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). This example suggests that some closures are driven by managers or owners who are unwilling to operate with a union due to a general dislike of unions. Second, consider Boeing's production of 787 airplanes. From 2011–2021, it moved the production of all 787s from a unionized plant in Washington to a non-union plant in South Carolina (Cameron, 2020). According to a Boeing executive, the motivation "was not the wages we're paying today. It was that we cannot afford to have a work stoppage, you know, every three years" (Greenhouse, 2011). Boeing's strategy illustrates how some firms can avoid working with unions, in this case by shifting production away from unionized establishments. Both examples show how unionization may cause establishment closures even if it does not lead to large wage or productivity effects.

This paper assesses whether these examples generalize by analyzing the effect of unionization on establishment employment and survival. We then test whether firms' ability to avoid working with new unions and managers' or owners' general opposition to unions help explain the overall negative effects of unionization we document. Our setting is around 27,000 U.S. private-sector union certification elections from 1981–2005 through the National Labor Relations Board (NLRB). To implement our analysis, we link data on union elections and contracts to administrative Census data on establishment employment, survival, and productivity.

We analyze these elections using a novel research design that extends standard difference-in-differences techniques with falsification tests from the regression discontinuity (RD) extrapolation literature. This strategy allows us to avoid biases from only comparing very close elections and to estimate treatment effects that include elections that won by larger margins of support. Using this strategy, we find that unionization decreases establishment employment, primarily by lowering their likelihood of survival. We estimate a five-year survival effect of four percentage points (pct. pts.) relative to a survival rate of 82% for establishments where the union lost. We also find bigger employment declines for larger margin-of-victory elections. Finally, we document significant effect

[1] Establishments are distinct locations where employees work (e.g., a manufacturing plant or retail store). Firms are groups of establishments under the same ownership. Unionization in the U.S. generally occurs at the establishment level.

[2] 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 summary of the wage and productivity literatures and whether it is consistent with these effects causing establishment closures.

heterogeneity across three broad industry groups: manufacturing, services, and other blue-collar and industrial sectors ("other" industries).[3] In the service sector, where most recent union organizing has occurred, we find small and sometimes insignificant effects of unionization. Alternatively, the overall employment and survival declines are driven by elections in manufacturing and the other industries. For example, the ten-year survival effect for manufacturing elections is eight pct. pts.

Next, we test whether firms' ability to avoid working with new unions or employer opposition to unions help explain the overall negative effects we document. For this analysis, we focus on manufacturing elections because we have better data to test specific parts of our hypotheses, and it is the largest sector where we find substantial negative employment effects.

Our first hypothesis is that some firms can avoid working with new unions by shifting production from a newly unionized establishment to their other establishments. To test this, we first estimate whether the effects of unionization are larger at establishments part of multi-establishment (or multi-unit, MU) firms than single-establishment (SU) firms. We find significantly larger employment and survival decreases at MU firms. For example, the ten-year survival effects are twelve pct. pts. versus three pct. pts. at MU and SU firms, respectively. This heterogeneity is consistent with MU firms avoiding working with 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 likely produce similar products to the election establishment, which makes shifting production to them the easiest. However, these effects are insignificant five years after an election, at which point the firm may have shifted production to *new* establishments. Both pieces of evidence support firms avoiding unions through production shifting as one explanation for the overall impact of unionization.

Our second hypothesis is that the effects of unionization are greater when managers or owners are more opposed to the union. To test this, we estimate treatment effect heterogeneity using two proxies for employers' opposition. First, we estimate effects separately for MU firms with and without any *other* unionized establishments. Survey evidence indicates that less unionized firms would more "vigorously resist dealing with unions," and that some non-unionized firms were motivated by a philosophical opposition to unions (Freedman, 1979; Foulkes, 1980). Additionally, similar to Selten (1978)'s "chain store paradox" a non-unionized firm might close a newly unionized establishment to convey its aggressive stance on unions, even if it would not be profitable to close when considered in isolation. Supporting our 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 employers' opposition to the union is the amount of delay time during the election process. Election delay time is a proxy for employers' opposition because it is a key way

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

that they attempt to influence elections. For example, in "Confessions of a Union Buster," 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 the date the union filed for the election and the election date. We estimate separate treatment effects for elections with shorter and longer delays, and find significantly larger employment and survival decreases following longer delay elections. For example, the ten-year survival effect for MU elections in the top tercile of election delay times is 20 pct. pts. versus 7 pct. pts. for the bottom tercile.

Finally, we test for effect heterogeneity by establishment productivity. If the survival decreases we document are driven by the conventional wage or productivity explanations, many theories of firm dynamics predict larger survival declines for lower-productivity establishments. However, we do not find significant differences in the survival effects of unionization between establishments with different baseline total-factor productivity, calculated from the Annual Survey of Manufacturers (ASM). This evidence is more consistent with our alternative explanations for why unionization leads to establishment closures than the conventional wage and productivity explanations.

Our overall employment and survival estimates present a puzzle relative to past research on union elections. Specifically, this research has not found that recent unionization led to large wage increases or productivity declines that would drive establishments out of business (see Appendix C). Our two alternative hypotheses provide resolutions to this puzzle. 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, our evidence on employer opposition suggests that the overall negative effects are driven by manager's or owner's dislike of working with unions rather than economic costs of unions. This interpretation is consistent with our finding of no treatment effect heterogeneity by establishment productivity.[4] However, we cannot rule out that our proxies for employer opposition simply capture rational expectations of the economic costs of unions. This interpretation is supported by research suggesting direct costs of union elections (e.g., Lee and Mas (2012)'s evidence of equity declines following unionization). Overall, while we do not measure the direct economic costs of unionization, our results suggest that the effects of unionization on employment and survival may substantially overstate those costs.

We next summarize our econometric methodology. We start by implementing a difference-in-differences (DiD) design that compares establishments where unions won versus lost. However, we only include a limited bandwidth of vote shares around the 50% threshold and implement multiple falsification tests of our identifying assumption within this bandwidth. Our identifying assumption is that outcomes at establishments with different election vote shares but the same baseline characteristics would have followed parallel trends had no election occurred. To support this assumption, we first show that only conditioning on baseline employment and industry yields similar pre-election employment and payroll growth rates between establishments with winning and

[4] As further support for this interpretation, survey evidence has not found that the firms most opposed to unions are also where unions were most likely to be the costliest (Freedman, 1979; Bronfenbrenner, 2001)

losing elections. This similarity holds when we add much richer baseline covariates and for up to ten years before elections. Furthermore, we show that our treatment effects are increasing in the share of workers in the bargaining unit and not driven by firm-wide trends.

Next, we assess additional testable 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). First, we show that the similarity in *pre-election* employment growth rates holds across all vote shares in our bandwidth. In other words, we test for "pre-trends" between finer vote-share groups than only winning and losing elections. 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 losing elections with different vote shares. These results show that our identifying assumption holds for subsets of observations where we observe untreated potential outcomes, supporting our (untestable) parallel trends assumption.

Our analysis combines features of regression discontinuity (RD) and panel data methods that have been used to analyze union elections.[5] Although RD methods 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 than elections with larger margins of support. To address these issues, we expand our bandwidth to include 20–80% vote-share elections and use the panel dimension to account for selection into winning versus losing elections. The wider bandwidth also gives us more power to estimate heterogeneous treatment effects. Relative to the other panel-data analyses, we better exploit observing vote shares to test our identifying assumption.[6] These tests can be implemented in other DiD analyses where the "forcing variable" is observed (e.g., Ganong and Noel (2020); Harju et al. (2021)).

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 negative survival effects.[7] Our findings also generalize other research finding employment declines following successful union elections in specific sectors (Sojourner et al., 2015; LaLonde et al., 1996).[8] 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

[5]See DiNardo and Lee (2004); Sojourner et al. (2015); Knepper (2020); Bradley et al. (2017) for RD analyses and Freeman and Kleiner (1990b); LaLonde et al. (1996); Lee and Mas (2012); Dube et al. (2016); Goncalves (2021) for panel data analyses. Frandsen (2021) also combines these methods by implementing a RD design on first-differenced outcomes.

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

[7]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 the manipulation around the threshold.

[8]LaLonde et al. (1996) analyze the employment and output effects of manufacturing union elections from 1977–1989 using a difference-in-differences design. We improve on their analysis on several dimensions. First, they do not analyze the effect on establishment closures, which makes interpreting the results conditional on survival difficult. Second, due to their smaller sample size, their pre-trend estimates are often imprecise, making it difficult to evaluate the parallel trends assumption.

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 vary substantially across sectors.

Our evidence supporting the union avoidance and employer opposition hypotheses is novel relative to prior economics research, but consistent with labor relations research. For example, Bronfenbrenner (2000, 2001) report similar results from a survey of union organizers in the 1990s. She finds survival effects 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' broader union avoidance tactics (Freeman and Kleiner, 1990a; Kleiner, 2001; Flanagan, 2007). In particular, this result complements Ferguson (2008)'s finding that successful elections with unfair labor practice charges (another proxy for employer opposition) are less likely to reach first contracts. Finally, our production shifting evidence is consistent with firms becoming less unionized during this time by investing in and opening non-union establishments (Kochan et al., 1986a; Verma, 1985).

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 union avoidance and employer opposition hypotheses. Sections 7 and 8 discuss our results and conclude.

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, when a union represents a group of workers, their employer is required to bargain with the union over the conditions of employment.[9] 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.[10] 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 NLRB election. The organizing drive is initiated by workers at the establishment, either on their own initiative or prompted by outreach from a union. The first step is getting cards signed indicating union support by workers in the "bargaining unit" (i.e., the workers the union would represent). A bargaining unit generally 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 disagreements over the bargaining unit composition,

[9] These include wage and non-wage compensation and promotion, grievance, and layoff policies (Slichter et al., 1960).

[10] For example, the currently debated *Protecting the Right to Organize (PRO) Act of 2021* would limit employers' ability to campaign against union elections and increase penalties for unfair labor practices during elections.

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).

Unions and employers often actively 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 or wearing pro-union attire), and enlisting the support of community groups (Bronfenbrenner and Juravich, 1998). Employers also use many campaign tactics, including holding one-on-one meetings with supervisors, requiring employees to attend "captive audience meetings," and hiring "union avoidance" consultants and law firms (Logan, 2002). Finally, although there are legal restrictions on firing pro-union workers and threatening to close establishments, these tactics still occur (Weiler, 1983; Schmitt and Zipperer, 2009).

If a majority of workers votes for the union, the union is certified by the NLRB to represent the bargaining unit. After certification, the employer is required to bargain "in good faith" with the union, but they are not required to reach an agreement.[11] If a contract is not reached one year after certification, the employees can vote out the union in a *decertification* election. The NLRA also establishes limitations on whether firms can close newly unionized establishments. Generally, establishment closures or relocations violate the NLRA if they are motivated by "anti-union sentiment." Instead, closures or relocations motivated by "economic reasons", do not violate the NLRA. In actual cases, the NLRB considers whether the firm's statements suggest an anti-union motivation, whether the firm was planning the closure before the election, and the timing of the closure relative to the election (Munger et al., 1988). Finally, closing an *entire* firm rather than a single establishment will generally not violate the NLRA regardless of the motivation. Consequently, it is legally easier to shut down after an election at a single-establishment firm than a multi-establishment firm.

NLRB elections are the primary method to gain union representation at a private-sector establishment. However, some unionization occurs without an NLRB election. First, the NLRA does not cover all workers (GAO, 2002).[12] Second, covered workers can also gain representation through voluntary, "card check" recognition. However, card check is much less common than elections.[13]

Selection into Union Elections and the Determinants of Winning 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 additions 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.

Workers, employers, and other factors could all influence whether the union wins an election. For our empirical strategy, the concern is that vote shares may be related to future establishment

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

[12] Some workers lack collective bargaining rights (e.g., some small business employees, independent contractors, domestic workers, and "agricultural laborers"). Others have collective bargaining rights but are not covered by the NLRA (e.g., airline and railroad workers covered by the *Railway Labor Act* and public-sector workers covered by federal, state, or local statutes).

[13] 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.

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 the union. This would generate a positive bias between vote shares and establishment growth. Alternatively, firms that expect to become more productive may campaign harder against unions, leading to a negative bias.

Research on election outcomes finds that all these factors play a role. The most consistent finding is higher union win rates for smaller bargaining units (Heneman and Sandver, 1983; Farber, 2001). Win rates also vary substantially across industries (Bronfenbrenner, 2002). These findings motivate our first specification that just conditions on 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 address the concern that firms' anti-union 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 employment growth. Further supporting our identifying assumption, Dube et al. (2016) find similar productivity pre-trends for nursing home elections, and Lee and Mas (2012) find similar stock-market trends, which is a stronger test since it incorporates expectations of future productivity growth.

Motivation for Getting Away from the RD Election Threshold An advantage of our empirical strategy is that it does not rely on only comparing elections around the 50% threshold. One motivation for this is non-random sorting of elections just around this threshold ("vote-share manipulation"). Figure 2 Panel A plots the vote-share distribution of elections in our sample with at least 50 votes and shows a large missing mass of elections that barely win, as documented in Frandsen (2017).[14] Frandsen (2021) shows that this causes large differences in observable establishment characteristics across the threshold (e.g., 13–22% employment differences).

Another motivation 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 2 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 close elections (e.g., the median (mean) for 51% vote-share elections is 118 (223) days versus 11 (57) for 60% elections). Since delays can dampen unions' bargaining power, this delay may lead to different effects for close elections than even slightly higher vote-share elections.

Additionally, for close elections, firms may delay bargaining anticipating union decertification. Figure 2 Panel C plots the probability that a certification election is followed by a *decertification*

[14] Given the discreteness of the running variable and the fact that our sample includes elections with a small number of votes, we cannot use the vote-share density figure for the *entire* sample to visually detect manipulation. Consequently, we restrict the sample in Figure 2 to elections with at least 50 votes. Frandsen (2017) provides evidence of manipulation in this setting 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, which illustrates the actual distribution of vote shares in our analysis.

election at the same establishment in the next five years. It 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. This suggests that higher margin-of-victory elections may be more likely to reach first contracts, leading to more workplace changes. A final reason for effect heterogeneity by vote share is that unions that win with more support may be able to more credibly threaten to strike. Figure 2 Panel D supports this by showing that for winning manufacturing election, where strikes were more common, the probability of a post-election works stoppage increases with the vote share. Panels B-D show that several 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 establishment-level data from the U.S. Census Bureau. These data are uniquely suited to study union elections. First, they contain the universe of *establishments*, the level at which most elections occur. Analysis of more aggregated firm-level data would attenuate the effects of unionization. Second, the Census constructs high-quality longitudinal establishment links that allow us to separate real establishment exits from spurious exits due to administrative reasons or ownership changes (Haltiwanger et al., 2013). These links are important for our analysis because survival is a key outcome.

NLRB Election Data We combine data from multiple sources to construct a comprehensive dataset of union elections from 1962 to 2018. Specifically, we combine data from Henry Farber, J.P. Ferguson, and Thomas Holmes and public NLRB data.[15] The data contain vote counts that we use to define treatment and employers' names and addresses that we use to match elections to Census data. Additionally, they include the election filing date, the actual election date, and the case-closing date. We define treatment time based on the election filing date because this is the earliest date we observe for each election. We also use these dates to define the number of days between filing and holding an election, a proxy for employers' opposition to the union (see Section 6).

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. We combine data from Thomas Holmes and the FMCS. The data include notices of *initial contract negotiations* (i.e., the first contract after an election) and *renegotiations or reopenings* 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 possible. We use these data to measure whether an establishment has other workers already covered by a CBA and whether the establishment's firm has other unionized establishments.

[15] For duplicates across datasets, we pick one observation for each NLRB case number (see Appendix D for details). Appendix Figure A2 shows that this yields a similar number of cases each year to the number of cases from the NLRB's annual reports.

Employment, Payroll, and Survival Data Our primary establishment-level outcomes are from the Longitudinal Business Database (LBD). It contains annual employment and payroll for the universe of non-farm, private-sector establishments from 1976–2015 (Jarmin and Miranda, 2002). Our employment measure is the total number of employees in March of each year. The payroll measure is employees' total "wages, tips, and other compensation" over the entire year. Consequently, we expect larger effects on "event-time zero" 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 the last year the establishment has non-zero employment. Finally, we use the Fort and Klimek (2016) 2012 NAICS codes to classify each establishment into consistent industries across the entire period.

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, the employment and payroll data are sometimes received at the more aggregated EIN level. These aggregate measures are initially 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 longer-run outcomes (e.g., five to ten years) since the Census receives establishment-level measures at least every five years (see Appendix D for details).

Sample Selection and Matching Datasets Before matching the elections to the Census data, we impose sample restrictions to focus on elections likely to shift establishments' union status. Appendix Table A1 shows how these restrictions affect the number of elections. First, we restrict to elections from 1981–2005. Since the LBD is available from 1976–2015, this always gives us a five-year pre-period and ten-year post-period. Second, we drop non-representation cases (e.g., decertification elections). Third, we drop contested elections (i.e., elections involving multiple unions). These are often "union raids" involving incumbent unions and consequently may only change *which* union represents the workers (Sandver and Ready, 1998). Fourth, we drop elections with fewer than six eligible voters to ensure that the election could lead to a non-trivial increase in union representation.

Next, we implement a name and address matching procedure to link each election to a unique establishment in the LBD (our strategy follows Kline et al. (2019)). We first calculate a weighted average of the *Soft TF-IDF* distance between employer names and the geographic distance between geocoded addresses for each election and LBD observation. 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. See Appendix D for details.

We further restrict the matched sample based on requirements of our empirical strategy. First, we only keep the first election at each establishment. Consequently, our estimates reflect the effects of the *first* election at an 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 before the election, we exclude observations where we cannot evaluate this for at least three years. To keep our sample the same across specifications, we also require that each observation have non-zero payroll and employment one year before the election. This results in a sample of approximately 27,000 elections (see Appendix Table A1). Finally, for much of our analysis, we restrict the sample to 20–80% vote-share elections. This decreases our sample to 19,000 elections. The motivation for this restriction is that some of the tests of our identifying assumption discussed in Section 4 fail for the extreme vote-share elections. However, we show that all our main results are robust to instead including a narrower 30–70% bandwidth.[16]

Table 1 presents summary statistics for winning and losing union elections in our sample. The estimates confirm the patterns of selection into winning elections described in Section 2. In particular, winning elections are at establishments that are, on average, smaller, less likely to be part of multi-establishment firms, and more likely to already have another unionized bargaining unit. The differences, however, are fairly small for payroll per worker and establishment age.

4 Empirical Strategy and Identifying Assumptions

Our research design combines standard difference-in-differences (DiD) techniques with tests of our identifying assumption from the regression discontinuity extrapolation literature. Our identifying assumption is a conditional parallel trends assumption between elections with different vote shares. Since we observe vote shares that determine treatment assignment, we can assess several testable implications of this assumption that are not possible in a standard DiD setting.

Potential Outcomes To fix ideas, consider 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^0_{it}. Its unionized potential outcome is $Y^E_{it}(V)$, which depends on its cohort E and vote share V. This allows for dynamic treatment effects and heterogeneous treatment effects by vote share, respectively. We assume *no anticipation* before the year of the election (i.e., $Y^E_{it}(V) = Y^0_{it}$ for all $t < E_i$). Observed outcomes are thus[18]

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

[16]Specifically, our vote-share heterogeneity estimates in Figures 5 and 6 show that the overall estimates are not driven by the 20–30 or 70–80% elections. Additionally, Tables A4, A5, and A6 present the heterogeneity estimates with a 30–70% bandwidth and show that the results are qualitatively the same although sometimes less precise than with the narrower bandwidth.

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

[18]Here, we assume that losing elections have no causal effect. However, everything would still go through if losing elections have a causal effect that is *constant across vote shares*, although this would modify the interpretation of our treatment effects.

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

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

where γ_i are establishment fixed effects (FEs) and α_t are year FEs.[19] 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*. Specifically, 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

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

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 only voting for effective unions). Second, the assumption is stronger than the standard DiD assumption because it requires parallel trends by vote share instead of only, on average, between the treated and control observations. Yet, this stronger assumption yields a richer set of testable implications. Additionally, we do not need to impose this assumption across the entire vote-share distribution. Instead, we can only assume it for some bandwidth around the 50% threshold and assess the subsequent testable implications within that bandwidth. This illustrates how our strategy allows us to get away from just comparing elections around the 50% threshold without needing to assume parallel trends between *all* winning and losing elections. Finally, as discussed in Section 2, vote shares may be influenced by workers, employers, and other factors that could lead to violations of this assumption. This possibility motivates our conditioning on specific baseline X_is and assessing multiple testable implications of this 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.

[19] 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 $t = E_i - 1$. For survival, it is unclear what time-invariant characteristic FEs would capture. We include establishment FEs for log outcomes.

Testable Implications of our Identifying Assumption Our identifying assumption yields several testable implications. Intuitively, we observe untreated potential outcomes, Y^0_{it}, 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**

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

This test nests the standard DiD pre-trends test between all winning versus losing elections. Moreover, we can test for similar pre-trends between finer vote-share groups. For example, we can estimate whether establishments where the union won by different margins of victory grew at different rates before the election by comparing pre-trend estimates for 50–60% versus 60–70% elections. This test mirrors the tests proposed by Angrist and Rokkanen (2015) and Bennett (2020) for regression discontinuity identification away from the threshold. They argue that conditional mean independence of untreated potential outcomes and the running variable for a given bandwidth around the RD threshold is strong support for being able to estimate treatment effects within that bandwidth. One reason that we only include 20–80% vote-share elections in our preferred specification is that for some outcomes, we find violations of equation 6 for extreme parts of the vote-share distribution.

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

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

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 losing elections with different vote shares.

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

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

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

This specification is the same as the single-cohort specification in 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).[20] Intuitively, our estimates come from only comparing winning and losing elections *within the same cohort* rather than making any "forbidden comparisons" between successful 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*.[21] 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.[22]

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.[23] 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). Next, 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).[24] We refer to this specification as the *pooled controls* specification. Finally, we interact all controls from the previous specification with cohort (i.e., year of election). This is our preferred *flexible controls* specification. The cohort interactions result in the analog to the "stacked regression" approach discussed previously. We show, however, that our results are robust to pooling controls across cohorts or only including the employment and industry controls.

[20] Gardner (2021) shows that the stacked regression approach, without any controls, estimates a convex weighted average of the cohort-specific treatment effects. Additionally, we test for negative weights on each cohort treatment effect in the exact specification we implement with controls using Sun and Abraham (2021)'s `eventstudyweights` package.

[21] With multiple cohorts, our identifying assumption is $\mathrm{E}\left[Y^0_{it} - Y^0_{it-1}|X_i, E_i, V_i\right] = \mathrm{E}\left[Y^0_{it} - Y^0_{it-1}|X_i, E_i\right]$. Thus, we do not require that selection into elections in the 1980s is the same as in the 2000s.

[22] This accounts for serially correlated outcomes across time and across elections at different *establishments* within the same *firm*. Our "stacked regression" approach to aggregate the δ_n estimates easily accommodates this clustering.

[23] 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, just 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.

[24] 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 controls in this specification with our three coarse industry groups (e.g., manufacturing, services, and "other"). This keeps them at the same level of granularity for our all industries and manufacturing estimates.

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 value of G_{it} could represent either an approximately 20% decline in intensive margin employment with no survival effects or a 10 percentage point decrease in the likelihood of survival. Since the growth rate accommodates exit, we can simultaneously evaluate pre-trends and interpret treatment effects, even 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.[25]

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 completely explained by a 10 pct. pt. decrease in survival). However, the residual part unexplained by exit could be 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 log outcomes are a useful complement to the DHS growth rate pre-trends.[26] 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 sometimes suggests intensive margin effects (e.g., large effects on log outcomes before large survival effects). Yet, we still recommend interpreting the treatment effects for log outcomes with caution since we do not completely address potential bias from selective survival.

Our parallel trends assumption in equation 5 imposes a specific functional form restriction for each of our outcomes (Kahn-Lang and Lang, 2020; Roth and Sant'Anna, 2021). First, we assume that log employment and payroll would have (conditionally) evolved in parallel, which we view as a reasonable restriction in this setting.[27] 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

[25] See Haltiwanger et al. (2013); Chodorow-Reich (2014) for general use and Arnold (2019); Davis et al. (2014) for DiD contexts. Often, the growth rate is defined annually, but we define it over longer time-horizons to measure cumulative changes.

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

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

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). For DHS growth rates, the outcome is approximately a linear combination of log changes and survival probabilities, so the previous two functional form assumptions imply parallel trends in the DHS growth rate.[28]

5 Results: Employment and Survival Effects

In this section, we estimate the effects of successful union elections on establishment employment and survival. We first analyze the differences in employment growth rates between establishments with winning and losing elections. Next, we implement several tests of our parallel trends identifying assumption described in Section 4. Finally, since we later focus on manufacturing, we present our estimates and falsification checks separately for elections in manufacturing.

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 elections with 20–80% vote shares. 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 (i.e., only year by cohort FEs), just baseline industry and employment controls, and the flexible control specification that interacts the controls by cohort (see Section 4).

Panels A. and B. show that establishments with successful elections had similar conditional pre-election growth rates to establishments with unsuccessful elections but experienced large relative employment decreases following the election. The first, "no control," estimates show that, without any controls, establishments where the union won had relatively slower pre-election employment growth rates than establishments where the union lost.[29] However, the next "industry + emp ctrls." estimates show that just conditioning on baseline employment and industry yields similar pre-election growth rates for DHS and log employment. As discussed in Section 4, we start with these two controls because they are prominent predictors of election outcomes and establishment employment growth. Starting one year after the election, this specification also shows decreased employment for establishments with successful elections. The effects stabilize approximately 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 the richer and more flexible controls. We later show that our "vote-share heterogeneity" tests of our identification assumption also yield very similar estimates with the "industry + emp" versus" flexible" control specifications.

To help interpret the magnitude and timing of the employment effects, Figure 3 Panel C additionally plots payroll and establishment survival estimates. Specifically, it includes estimates of

[28] Formally, $\mathrm{E}[\Delta \ln Y^0_{it}|X_i, V_i] = \mathrm{E}[\Delta \ln Y^0_{it}|X_i]$ 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 $\mathrm{E}[G^0_{it}|X_i, V_i] \approx \mathrm{E}[G^0_{it}|X_i]$.

[29] Without controls, the DHS employment pre-trends combine intensive and extensive margin changes, while log employment only captures the intensive margin. The measures are approximately the same when we control for establishment age.

DHS employment and payroll growth and establishment survival with the flexible control specification. We find that payroll initially declines faster than employment. This difference could be due to either compositional shifts to lower-wage workers or differences in the timing of the payroll versus employment variables described in Section 3. Five years after a successful election, the cumulative DHS employment and payroll growth rates are -0.13 and -0.14 lower, respectively, than establishments with unsuccessful elections (consistent with a 14% decrease in payroll or a seven pct. pt. decrease in survival). Appendix Figure A3 presents estimates from the same specification for *log* employment and payroll. These estimates allow us to reject five-year, pre-election growth rate differences of more than 3.5% for employment and 1.8% for payroll. In other words, we can rule out that establishments with winning elections grew more than 3.5% faster than ones with losing elections 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 from a lower likelihood of 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 illustrates how much of the DHS effect can be mechanically explained by the survival effect (see Section 4). The estimates show that five years after an election, establishments with successful elections are four pct. pts. less likely to survive, and this effect increases to five pct. pts. after ten years. Consequently, about two-thirds of the -0.13 five-year DHS employment growth rate estimate is mechanically due to decreased establishment survival. Finally, the relatively slower timing of the survival versus employment effects could be due to an increased risk of violating the NLRA when immediately closing an establishment following an election (see Section 2 for details).

Given our later focus on manufacturing, Figure 4 presents the same estimates including only manufacturing establishments. For these elections, we find similar pre-election employment growth rates (i.e., a lack of pre-trends) even with no controls. For example, Panel B shows that, without any controls, we can rule out five-year employment growth rate differences of more than five percent. One explanation for the lack of detectable pre-trends without controls is that by only comparing elections in manufacturing, we account for sector differences that the controls capture when we pool all industries. Additionally, for manufacturing, the magnitude of the treatment effects is 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 relatively small.

Vote-Share Heterogeneity Tests of our Identifying Assumption

Next, we provide more evidence that our results are driven by unionization by assessing several testable implications of our identifying assumption. We first present visual evidence of how treatment effects and pre-election trends vary across the vote-share distribution and then implement parametric versions of these tests. This also allows us to estimate treatment effect heterogeneity by vote share.

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

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

$\mathcal{V}^g$ are exhaustive subsets of the vote-share distribution.[30] 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 that lost by exactly 50% and discuss several reasons to expect this.

Figure 5 presents estimates from equation 10 for all manufacturing elections. The estimates include our *flexible controls* specification (see the following parametric test results for robustness to alternative controls). Panel A includes estimates for each vote-share group with DHS employment growth as the outcome. First, the five-, three-, and two-year pre-trend estimates are similar across almost the entire vote-share distribution relative to the omitted 20–30% elections (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. These results provide reassurance against the concern that our main estimates are driven by future productivity shocks correlated with vote shares. In that case, we would also expect these shocks to cause different outcomes for losing elections with different vote shares. Finally, the five- and ten-year treatment effect estimates for winning elections increase in the union vote share (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 still test our parallel trends assumption for survival by estimating whether losing elections' with different vote shares had different post-election survival rates. Reassuringly, the survival rates for all losing election vote-share groups are not statistically different than for 20–30% elections. For winning elections, however, the figure shows larger long-run survival declines for higher vote-share elections, although the group differences are not statistically significant.

Figure 6 presents vote-share heterogeneity estimates pooling together all industries. Panel A shows that for all 20–80% vote-share groups, we find very similar pre-election DHS employment

[30] We omit the establishment FEs because we only estimate this specification for outcomes where we do not include these FEs.

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 votes, and there are multiple reasons that 50% elections are quite different than elections where the union lost by slightly larger margins.[31] To show this, we estimate the 40–50% effects excluding 50% elections and 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 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 between any of the losing election groups with 20–50% vote shares. Furthermore, Panel B of Figure 6 shows that there is no evidence of differential survival rates between 20–30, 30–40, and 40–50% losing elections, even including the 50% elections.

Parametric Vote-Share Tests To complement the previous *nonparametric* analysis, we estimate a series of *parametric* vote-share heterogeneity 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 extensions. 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 only the *employment and industry* controls and with the *flexible* controls.

To implement these tests, we estimate a modified version of the specification in equation 8. Specifically, instead of interacting event-time with the winning indicator (e.g., $\mathbb{1}[V_i > .5]$), we include the following interactions with event-time[32]

$$\underbrace{\mathbb{1}[t - E_i = n]}_{\text{Event Time}} \times \begin{cases} \rho \cdot V_i & \text{if } n < 0 \\ \eta \cdot \mathbb{1}[V_i > .5] + \theta \cdot V_i + \tau \cdot [V_i - .5]^+ & \text{if } n \geq 0 \end{cases} \tag{11}$$

For the pre-period, ρ estimates a linear trend in employment growth rates by vote share for all 20–80% elections. For the post-period, we include an interaction with treatment so η captures the treatment effect for close elections (e.g., a linear RD estimate). Thus, θ estimates a linear trend in post-election outcomes for losing elections and $\theta + \tau$ captures this trend for winning elections.

[31] 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 exactly 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).

[32] $[V_i - .5]^+$ is equal to $[V_i - .5] \times \mathbb{1}[V_i > .5]$. Since we only estimate this specification for elections with 20–80% vote shares, we shift the vote-share variables to all start at zero (e.g., subtracting 0.2 from the V_i variables and 0.3 from the winning vote-share variable). This ensures that the vote-share coefficients only capture slope and not level differences.

Table 2 includes estimates of pre-election growth rate trends, ρ, for one to five years before the election. The estimates are for the 20–80% vote-share sample. We present separate estimates for all industries and manufacturing and for the *employment and industry* controls and *flexible* controls. Across all estimates, we never find significant pre-election growth rate trends.[33] This complements the previous nonparametric estimates by showing that the lack of pre-trends across the vote-share distribution holds formally testing for linear trends and only including more limited controls.

Table 3 presents estimates that test for post-election vote-share trends. We present estimates of separate slopes for losing elections (i.e., θ) and winning elections (i.e., $\theta + \tau$). This table includes our preferred *flexible control* specification, but Appendix Table A2 shows qualitatively similar results only including the *employment and industry* controls. Motivated by the previous issues with exactly 50% vote-share elections, we present estimates with and without excluding these elections.

The results for all industries in Table 3 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 do not detect significant trends for losing elections. The significant negative trends by vote share for winning elections indicate increasing treatment effects for larger margin-of-support elections. For example, we estimate a vote-share trend of -0.066 (SE 0.122) for losing elections and -0.389 (SE 0.149) for winning elections for five-year growth rates. For establishment survival, we never find significant trends for winning or losing elections (sometimes, the losing election trends are actually positive). The manufacturing estimates in Table 3 Panel B are similar to the estimates for all industries. Without excluding the 50% elections, we find negative although insignificant DHS employment trends for losing elections. However, dropping the 50% elections results in smaller trends for losing elections and large although insignificant vote-share heterogeneity estimates for winning elections (e.g., five-year DHS trend estimates of -0.072 (SE 0.199) for losing elections and -0.406 (SE 0.299) for winning elections). The survival estimates are also never significant for winning or losing elections.

Overall, these estimates confirm the nonparametric post-election estimates in Figures 5 and 6. First, they show that, excluding the 50% elections, the lack of vote-share trends in post-election employment growth for losing elections holds when testing for linear trends and only including the employment and industry controls. For establishment survival, we also cannot detect trends with and without excluding the 50% elections. We note, however, that the 95% confidence intervals sometimes include relatively large post-election growth rate differences for losing elections. Additionally, the estimates for winning elections provide a formal test of treatment effect heterogeneity by vote share. Specifically, for the overall DHS employment growth rate estimates, we find significant vote-share heterogeneity.[34] For survival, however, we do not find evidence of treatment-effect heterogeneity.

[33] A reasonable benchmark to assess the magnitude of the estimates is what they imply for the differences between 20–30 and 70–80% elections in Figure 6. The largest 0.05 coefficient we estimate, implies a difference in pre-election employment growth rates of around 2.5% between 20–30 and 70–80% elections. For all industries, the confidence intervals also allow us to reject large trends in pre-election growth rates (e.g., with flexible controls, we can rule out five-year growth rate differences of more than four percent). For manufacturing, however, the 95% confidence intervals would sometimes include relatively large differences.

[34] For manufacturing, the estimates are only significant at the 10% level although, Appendix Table A2 shows that when only including the industry and employment controls, we find significant estimates at the 5% level.

Additional Robustness Checks Next, we present two additional checks that further support our identifying assumptions. First, we assess whether our estimated effects increase in the size of the bargaining unit (Lee and Mas (2012) conduct a similar test). The motivation is that the relative share of unionized workers should mediate 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 their expectations of future company performance or managers' competence). Appendix Table A3 presents the coefficient estimates from interacting the three-, five-, and ten-year treatment indicators with the share of each establishment's total employment in the bargaining unit (see Appendix D for details). It shows that the three and five-year treatment effects are significantly increasing in the bargaining unit share for both outcomes. This is consistent with the effect being mediated by the share of workers gaining union representation. The interactions, however, are not significantly different than zero at the ten-year horizon. One explanation for the lack of persistence is that the bargaining unit share could change substantially in the 10 years following the election.

Second, Appendix Figure A4 plots DHS employment growth rate estimates with *ten-year* pre- and post-periods.[35] First, it shows no evidence of large pre-trends in employment growth rates up to ten years before elections for the all industry or manufacturing samples. Although we find significant six, seven, and eight year pre-period estimates in all industries, the estimates are economically small (e.g., approximately 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 approximately 3.2%. Second, the post-election effects are relatively stable between three to ten years after the election. For manufacturing, however, there is a slight increase in the effect from years five to ten.

Industry-Specific Employment and Survival 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., Figure 2 shows that 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, multi-establishment 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) or tradable industries with ties to their local area (e.g., hotels).

To estimate this heterogeneity, we classify elections into three industry groups: manufacturing, services, and other blue-collar and industrial sectors.[36] More than 70% of voters in service-sector

[35] 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). This is why we use the -5 to 10-year estimates for the main analysis.

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

elections are in healthcare (e.g., hospitals and nursing homes), security, restaurants, grocery stores, universities, and print media establishments. The other category includes agriculture, construction, mining, transportation and warehousing, utilities, and wholesale trade.

To estimate the industry heterogeneity, we use the following specification for a categorical heterogeneity variable H_i (e.g., the three industry groups)[37]

$$Y_{it} = \alpha_{t,E_i} + \sum_h \sum_n \delta_{h,n} \cdot \mathbb{1}[t - E_i = n] \times \mathbb{1}[V_i > .5] \times \mathbb{1}[H_i = h] + X_i' \beta_{n,E_i} + \varepsilon_{it}. \tag{12}$$

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

Table 4 presents the DHS employment growth and survival effects estimated separately for each industry group. First, there is limited evidence of employment growth pre-trends for any industry. The only marginally significant pre-period estimate is for the service sector, where we find the smallest main effects. Second, the overall employment and survival decreases are driven by 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 declines of more than four pct. pts.

Appendix Table A4 shows that the smaller effects of unionization in the service sector are robust to alternative sets of controls and sample selection criteria. Specifically, it presents the point estimate and standard error of the difference between the manufacturing and service-sector coefficients over each time horizon. The effects in manufacturing remain significantly larger when we (1) pool controls across cohorts, (2) use a 30–70% vote-share bandwidth, and (3) restrict the sample to elections where the size of the bargaining unit was at least 25% of total establishment employment. The last restriction shows that the smaller effects in the service sector are not because those elections are more likely to have a very small share of workers in the bargaining unit.

6 Results: Union Avoidance and Employer Opposition

After documenting the large overall impacts of unionization on establishment employment and survival, we explore two new hypotheses for these effects. Specifically, we test whether these effects are due to firms' ability to avoid working with new unions by reallocating production across establishments or due to managers or owners' dislike of working with unions. The motivation for exploring these hypotheses is that the conventional explanations for establishments closing after unionization have not been supported by prior research. Specifically, prior research has not found

---

workers (e.g., grocery store workers) are commonly referred to as part of service-sector unionization.

[37]This specification has two advantages relative to restricting the sample for each value of H_i, First, we can pool the control coefficients across heterogeneity groups and use all the data to estimate their coefficients. For all heterogeneity estimates, we also add the specific heterogeneity group as an additional control in X_i so that we account for any differential trends by the specific heterogeneity groups. Second, it allows us to easily conduct Wald tests of equality across the different heterogeneity groups.

that unionization in the past forty years led to large wage increases or productivity declines that could drive the firms out of business (see Appendix C for details). Furthermore, it would be puzzling for a union to make demands that push firms out of business, directly harming the union's members (Friedman, 1951). Both of our alternative explanations can rationalize large employment and survival effects even if unionization has a relatively small effect on wages and productivity.

For this analysis, we focus on elections in manufacturing for three reasons. First, the ways that firms can avoid working with new unions differ across sectors. In manufacturing, a common union avoidance tactic was shifting production from a firm's unionized establishments to its non-unionized establishments (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 working with new unions by opening a new, non-unionized firm that did the previous work of the unionized firm (Evans and Lewis, 1989). We focus on manufacturing because we can use our data's establishment to firm linkages to test for union avoidance via production shifting. Second, manufacturing is the largest sector where we find negative effects.[38] Finally, we only have detailed measures of establishment-level productivity in manufacturing that allow us to estimate effect heterogeneity by baseline productivity.

Union Avoidance via Production Shifting

Our first hypothesis is that firms avoid working with new unions by shifting production from a newly unionized establishment to their other establishments. This idea goes back to at least Ulman (1955a) who describes the difficulty in unionizing multi-establishment (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 consistent with evidence from this period that firms shifted employment and investment from their unionized establishments to non-unionized establishments (Kochan et al., 1986a; Verma, 1985).

Single- versus Multi-Establishment Firm Heterogeneity Since production shifting is only possible for firms with multiple establishments, we start by estimating whether the effects of unionization are larger at establishments part of multi-establishment (or multi-unit, MU) firms versus single-establishment (SU) firms. Specifically, we define "an election at a MU firm" based on whether the firm had at least two establishments under its control one year before the election.

Figure 7 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

[38]For all sectors in our manufacturing and other industry groups, manufacturing makes up 54% of elections compared to 18% for transportation and warehousing (the next largest sector). Weighting by eligible voters, it comprises 69% of voters.

negative and significant. For the establishment survival estimates in the right panel, the differences are even more striking. For all time horizons, the effects are significantly larger for MUs, and none of the 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 A5 shows the robustness of these estimates to (1) including controls pooled across cohorts and (2) instead using 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 "pooled controls." For the 30–70% bandwidth, we still estimate substantially larger survival effects for MU firms (e.g., six pct. pts. at the ten-year horizon), but the larger standard errors only lead to a significant difference at the five-year horizon.

We interpret these results as showing that the effects of unionization on establishment survival in manufacturing are driven by establishment closings at MU firms. For the overall employment declines, the effects are also significantly larger at the multi-establishment firms but still significant for SUs. This evidence is consistent with MU firms responding to unionization by shifting production across establishments, which we investigate more directly next.

Estimating Employment Shifting after Union Elections Next, we directly test whether manufacturing firms avoid working with new unions by shifting production to other establishments. Specifically, we analyze whether a successful election at one of a firm's establishments 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 (Boehm et al., 2019; Giroud and Mueller, 2017). 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.[39]

To construct the sample for this analysis, we start with all manufacturing elections in one year 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.[40] We then calculate these establishments' DHS employment growth rates relative to one year before the election. Finally, we stack these observations from all cohorts and estimate a modified version of our main specification 8.[41]

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

[40] 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 downward bias our overall spillover estimates.

[41] 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

The two differences from our main specification are that (1) relative time and vote-share variables are defined from the election at the firms' *other* establishment, and (2) we weight the regression by each establishment's share of its firms' total employment.[42] 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 firm and establishment.[43]

Figure 8 Panel A plots the employment effects of a successful election on the firms' other establishments. It presents estimates that include all manufacturing establishments and that only include establishments in the same three-digit NAICS industry as the election establishment. For all establishments, there is no evidence of relatively higher employment growth at the other establishments following successful elections. There are two things to note about this result. First, even if firms engaged in production shifting, 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 production shifting more costly. Second, it is reassuring that we do not estimate *lower* post-election employment at the *other* establishments of firms with successful elections. If establishment-level productivity shocks bias our estimates of the direct effects of unionization, we might expect some of these shocks to be firm-wide. Yet, the estimates in Figure 8 allow us to rule out differences in five-year DHS employment growth rates of more than -0.04 which is much smaller than our overall estimate for elections at MUs of -0.21.

We find significant employment growth effects when we restrict the sample to other establishments that produced similar products to the election establishments. The solid estimates in Panel A plot the estimates from just the firm's other establishments in the same three-digit NAICS industry as the election establishment. Two years after the election, we estimate a 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. However, it becomes insignificant five years after the election and remains insignificant after ten years. Additionally, Table 5 includes survival estimates and shows that some of the increased employment growth is due to an increased likelihood of survival.

Figure 8 Panel B splits up the same-industry elections based on whether the *election establishment* made up a large share of the firm's total employment. The motivation is that we would not expect to have enough power to detect spillovers when the election establishment was only a small share of the firm's overall employment. Specifically, we split up elections based on whether the election establishment was more than 10% of the firm's employment in the same three-digit NAICS industry during the election year. The estimates in Panel B show that the overall increase in other establishments' employment growth is driven by relatively large elections. This is reassuring because these are the elections where we would expect to detect the most production shifting, but

establishments can switch to different `firm IDs`. This also motivates our two-way clustering by firm and establishment.

[42] For the denominator, we only include employment at establishments in the sample so the employment weights sum to one.

[43] We use each establishment's firmid during the election year (e.g., the clustering variable is fixed over time). Yet, since an establishment can appear multiple times in the sample, establishment clustering is not nested by firmid clustering.

we would not expect potential threats to our parallel trend assumption to be more pronounced.

The magnitude of these employment shifting estimates is also economically significant. For the same-industry estimates, the DHS employment growth rate estimates of around 0.04 are consistent with a two percentage point increase in the survival probability of the firms' other establishments. When we focus on high-employment share elections, the effects are even larger between 0.07–0.09. As a benchmark, the direct three-year DHS employment growth rate effect of unionization for elections at MU manufacturing establishments 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. First, our spillover estimates are average establishment-level employment changes, while we would need firm-level estimates to calculate the total share offset by reallocation.[44] Second, we focus on a specific subset of establishments where we are most likely to detect spillovers. However, calculating the total share offset requires the total firm-level employment changes (e.g., the estimates for all manufacturing establishments where we do not have enough power to rule out significant spillovers).

Overall, this evidence of successful union elections leading to faster employment growth at the firm's other establishments is consistent with firms shifting production away from newly unionized establishments. Additionally, the survival increases 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 does not include shifting production by opening *new* establishments or to establishments in other countries (see Bluestone and Harrison (1982) and Bronfenbrenner (2000) for evidence of shifting production internationally following successful union elections).

Manager and Owner Opposition to Unions

Our second hypothesis is that unionization leads to more adverse effects when firms' managers or owners are more opposed to unions.[45] In other words, some of the employment and survival effects we document may be driven by anti-union animus rather than direct economic costs of unions. 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, the possibility that anti-union animus may motivate firms' responses to unionization is

[44] We conduct an establishment-level analysis for two reasons. First, the longitudinal establishment 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 age, baseline employment, and time-varying industry controls that explain some of the employment growth variation.

[45] We cannot distinguish between whether the opposition is driven by the preferences of *managers* or *owners* (Jensen and Meckling, 1976). Subsequently, we refer to this hypothesis as "manager" or "employer" opposition. However, in the U.S., a large share of owners are involved in management so we do not need to make the distinction for these firms (Kim et al., 2022).

central to U.S. labor law.[46] To test this, we estimate treatment effect heterogeneity based on two proxies for managers' opposition. We show that these proxies have more power than conventional measures like baseline productivity for predicting where unions cause to the most adverse effects.

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 from this period that non-unionized firms (e.g., firms with no unionized establishments) were more opposed to unions than (partially) unionized firms. For example, Freedman (1979) and Kochan et al. (1986b) show that less unionized firms 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 up our elections at MUs based on whether we observe an FMCS contract at any of the firm's establishments in the five years before the election (see Appendix D for details). Since the contract data start in 1984, we classify unionized versus non-unionized firms starting in 1985 and show robustness to instead starting in 1990. Figure 9 presents the DHS employment growth and survival estimates for elections at unionized versus non-unionized firms. The estimates are presented the same as the heterogeneity results in Figure 7. For overall employment growth rates, elections at non-unionized firms lead to larger employment decreases than elections at unionized firms. These differences are significant at the five- and ten-year horizons. For establishment survival, the differences are rather 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).

Appendix Table A6 shows that the larger effects at non-unionized firms are robust to alternative sets of controls and sample selection criteria. Specifically, it presents estimates of the difference between effects at unionized versus non-unionized firms when (1) pooling controls across cohorts, (2) classify unionized versus non-unionized firms starting in 1990, and (3) using 30–70% vote-share bandwidth. The differences between estimates for non-unionized versus unionized firms are larger than our baseline specification when we define firms' unionization status starting in 1990. For the other two specifications, the estimates are qualitatively the same as our baseline estimates.

These estimates show that the long-run negative effects of unionization are substantially larger at entirely non-unionized firms. This evidence is consistent with these firms being more opposed to and more rigorously resisting unionization. An alternative explanation for this result is that MU firms respond excessively to the first successful election 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 its aggressive stance on unions, even if it would not

[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; Garwin Corp, 1968; Wright Line, 1980; Weather Tamer. v. NLRB, 1982)

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

be profitable to close that establishment when considering it in isolation.[48] Both explanations imply that the employment and survival effects of unionization may be much larger than implied by the direct economic costs of operating a unionized establishment.

Election Delay Time Our second proxy for managers' opposition to the union is delay during the election process. The motivation is that employers can use tactics that delay the election date to try to win the election. First, delay itself can reduce support for the union. In "Confessions of a Union Buster," Levitt and Conrow (1993) write that the NLRA "presents endless possibilities for delays, roadblocks, and maneuvers that can undermine a union's efforts and frustrate would-be members" and that delay "steals momentum from a union-organizing drive." Additionally, other tactics employers use to influence elections also cause delay (e.g., challenging the composition of the bargaining unit or filing unfair labor practices). Furthermore, delay time is associated with lower election win rates which supports is being a proxy for the intensity of employers' 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 for details). The average delay in our sample is 62 days, and the 10th and 90th percentiles are 31 and 80 days. Appendix Figure A5 shows that our delay time measure 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 also hold conditioning on election characteristics that may be correlated with delay time.

To analyze whether the effects of unionization differ by delay time, we first estimate treatment effects separately by terciles of the within-year delay time distribution. Figure 10 plots the estimated effects for the first and third terciles for DHS employment growth (left panels) and survival (right panels). Panel A includes all elections and Panel B just includes elections at MU firms. The p-values are from testing whether the first and third tercile estimates are equal. Across both figures, the effects of unionization on employment and survival are larger for elections in the top tercile of the delay time distribution. For elections at MUs, the first versus third tercile estimates are significantly different for both outcomes at the three- and ten-year horizons (e.g., the ten-year survival effects for the top and bottom terciles are -0.20 (SE 0.037) and -0.071 (SE 0.036), respectively).

Next, we assess the robustness of these results to instead using a continuous measure of delay time. Although the tercile specification is appealing because it only relies on within-year variation and allows for a flexible, functional form, we may have more power using continuous delay time. To implement this, we add an interaction between the event-time treatment indicators and the *log* delay time to the specification in equation 8.[49] Table 6 presents the coefficients on the log election delay 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. For

[48]As an example, Walmart switched to pre-packaged meat across *all* stores days after ten meat cutters at one Texas Walmart unionized in 2000 (Zimmerman, 2000).

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

the ten-year survival effect, an approximately 10% increase in delay time is associated with a .7 pct. pt. decrease in the survival probability.[50] Columns (3) and (4) show that the effects are robust to including the controls pooled across cohorts. Columns (5) and (6) address the concern that the election delay time measure simply captures larger bargaining units. These columns show that our estimates are qualitatively the same when we first residualize log delay time on bargaining unit size deciles, although the ten-year estimates are only significant at the 10% level.

Our interpretation of these results is that the negative effects of unionization are largest when the employer initially campaigned harder against the union. This is supported by anecdotal accounts linking election delays to the intensity of firms' anti-union campaigns (Levitt and Conrow, 1993). Another interpretation is that delay may be a proxy for hostile labor relations. For example, more adversarial unions and management might have more disagreement before the election that could delay the process. Overall, this heterogeneity adds to our results showing that managers' opposition to unions plays a role in the overall negative effects of successful union elections.

Baseline Productivity Heterogeneity Finally, we estimate whether unionization leads to larger survival decreases at more or less productive establishments. If the survival decreases we document are driven by the conventional wage or productivity explanations, many theories of firm dynamics predict that the survival declines should be larger for lower-productivity establishments.[51] Consequently, larger survival declines at less productive establishments would be consistent with these effects being driven by unions increasing wages or lowering productivity. Alternatively, our hypotheses predict that the negative survival effects will be the largest at establishments with the most opposition to or ability to avoid unionization, which may not be the least productive.

To measure establishment TFP for our manufacturing elections, we use cost-share productivity measures from the Annual Survey of Manufacturers (ASM) and Census of Manufacturers (CM) calculated by Foster et al. (2016). We use within-industry TFP comparisons to address measurement or productivity differences across industries. Specifically, we classify each establishment into a productivity terciles based on its pre-election, within year and six-digit NAICS industry TFP ranking (see Appendix D for details). Figure 11 plots the estimated effects for the first and third terciles of the TFP distribution. We find evidence that the three- and five-year employment and survival effects are larger for lower-productivity establishments. But, these differences are never significant and, at the five-year and ten-year horizon, are not economically very large (e.g., -0.066 (SE 0.023) versus -0.041 (SE 0.022) at the five-year horizon). Appendix Figure A6 shows that these patterns hold when we separately estimate this heterogeneity for only MU firms.

Overall, we do not interpret this evidence as supporting economically larger survival effects

[50] 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.

[51] For example, in Hopenhayn (1992), firms exit when their productivity declines below an exit threshold, implying that lower productivity firms are more likely to exit in response to productivity declines. Similarly, (Berger et al., 2021) show that minimum wage increases cause relatively larger employment declines at less productive firms in oligopsonistic labor markets with firm heterogeneity. Consistent with these predictions, empirical research finds that minimum wage increases and trade liberalization disproportionately decrease survival for less productive establishments (Dustmann et al., 2022; Pavcnik, 2002).

for less productive establishments. Thus, the evidence is more consistent with our alternative explanations for why unionization leads to establishment closures than the conventional wage and productivity explanations. Additionally, the evidence does not support the "Swedish theory" of centralized bargaining where collective bargaining reallocates employment to more productive firms by driving low-productivity firms out of business (Edin and Topel, 1997).

7 Discussion: New Explanations for Unions' Effects

Employment and Survival Effects of Unionization

We first show that successful union elections substantially decrease establishment employment and survival, especially in manufacturing and other blue-collar and industrial sectors. Relative to past research on this question, our novel empirical strategy allows us to avoid issues from only comparing very close elections, while also better validating our identify assumptions than simply comparing all winning versus losing elections. The most comparable results to ours are Frandsen (2021)'s regression discontinuity estimates also using the LBD. We qualitatively match his short-run employment and long-run survival declines but find somewhat smaller effects (e.g., five-year survival effects of 4 versus 8–10 pct. pts.). Some explanations for this are different samples or empirical strategies.[52] Additionally, our smaller employment effects and insignificant 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 the null effects that DiNardo and Lee (2004) found for establishment survival and employment. A potential explanation is that the LBD longitudinal linkages we use to define survival are of higher quality than linkages in the telephone-book-based InfoUSA or the LRD data used by DiNardo and Lee (2004).[53] Finally, although long hypothesized (e.g., the exchange between Friedman (1951) and Ulman (1955b)), we provide the first evidence that the effects of unionization vary substantially across sectors.

The effects of unionization on employment and survival are relevant for workers voting in elections and organizers planning elections. For example, American Compass (2021) and Cohen and Hurd (1998) find that workers cite fear of "weakening [the] company" and "job security" as reasons why they would not vote to unionize. Bronfenbrenner (2000) also provides examples of firms claiming that unionization would force them to close. These concerns have gained renewed prominence during the recent resurgence of organizing activities (the SBWorkers United union claimed that Starbucks closed unionized stores in retaliation for organizing activities (Russ, 2022)). Our results show that these concerns were warranted but only in some sectors. Consistent with workers being less likely to vote to unionize when threats of establishment closure were more credible, from 1980 to 2010, the election win rate increased substantially in services relative to manufacturing.[54]

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

[53] See, Jarmin and Miranda (2002) and Crane and Decker (2019) for comparisons of the linkage quality across these datasets.

[54] From 1980 to 2010, the win rate of elections in the service sector increased from around 50% to above 60% while the win rate for elections in manufacturing remained below 40% for most of the period.

Implications of Union Avoidance and Employer Opposition

The conventional models of how unionization could decrease employment and survival operate through a standard establishment-level supply and demand framework. First, unions could decrease employment by raising wages and moving the establishment down its labor demand curve (Nickell and Andrews, 1983). Alternatively, unions could lower productivity and shift down the establishment's labor demand curve (Brown and Medoff, 1978). In either case, large employment decreases imply either large wage increases, productivity declines, or very elastic labor demand, which is inconsistent with prior estimates. Consequently, our estimates present a puzzle for these models. The evidence in the second half of the paper on union avoidance and manager opposition provides alternative reasons for large survival and employment effects even without large wage or productivity effects.

First, we provide evidence consistent with some of the effects of unionization being driven by firms reallocating production away from newly unionized establishments. This evidence helps resolve the previous puzzle because small wage or productivity effects could lead to large establishment-level survival decreases if firms can cheaply shift production across establishments. Although we are unaware of estimates of how costly such shifting is, there is evidence that firms designed their production networks to make it easier to reallocate production in response to unionization.[55] Additionally, although we focus on production shifting in manufacturing, firms in other sectors had different ways of avoiding working with new unions without reducing production. For example, Hatton (2014) documents firms replacing unionized workers with independent contractors in several industries, and Evans and Lewis (1989) document construction firm owners opening separate non-union firms to avoid hiring unionized workers.

Second, we show that the negative effects of unionization were largest when the managers or owners were likely more opposed to the union. One interpretation of these results is that employers' opposition and attempts to avoid unions were driven by their dislike of working with unions. This interpretation resolves the previous puzzle because idiosyncratic dislike rather than economic factors caused the negative effects. It is also consistent with Foulkes (1980) who documents that some non-unionized firms were motivated by philosophical opposition to unions and 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."

On the other hand, we cannot rule out that employers' opposition was completely driven by rational expectations of the economic costs of unions. This interpretation is supported by evidence suggesting direct costs of unions (e.g., stock price declines following successful elections and a literature on unions reducing profits (Lee and Mas, 2012; Freeman and Medoff, 1984)). However, even part of Lee and Mas (2012)'s results do not support the interpretation that our estimates are completely driven by direct costs of unions. In particular, they do not find stock price declines for close union elections, whereas we find negative effects for these elections (Frandsen (2021) also notes

[55] Bluestone and Harrison (1982) described 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.

this puzzle). One way to reconcile these results is that our negative estimates for close elections may be driven by our alternative hypotheses while economic costs may also play a role for larger margin-of-victory elections. Another piece of evidence inconsistent with the rational expectation interpretation is survey evidence showing that the firms most opposed to unions were not those who also expected unions to be the costliest. For example, Freedman (1979) finds that non-unionized firms placed the most weight on resisting unions, but these firms also expected unions to be the least able to bargain for higher wages. Overall, since we do not estimate the direct costs of unions, we cannot rule out either interpretation. Yet, our evidence shows that further understanding the cause of employers' opposition to unions is crucial for fully understanding their effects.

The role that production shifting and anti-union animus 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., firms or sectors) would result in more muted negative effects than we document.

Our estimates of effect heterogeneity across sectors add to a recent body of 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) found negligible employment effects of minimum wages in the service sector but larger negative effects in manufacturing. The production shifting channel we document may also explain why these other policies 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).

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 estimate treatment effects that include larger margin-of-victory elections. Our strategy and identifying assumption tests can be useful in other panel-data settings where the "forcing variable" is observed. Using this strategy, we show that unionization decreases establishment employment and likelihood

[56] Conventional measures of tradable industries (e.g., Mian and Sufi (2014)) do not necessarily identify which industries can shift production across regions. For example, hotels and mines are tradable products but are relatively immobile due to fixed local factors of production. Consequently, the production shifting channel we explore may not apply to all tradable industries

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 their other establishments. We support this by showing that the largest effects are at multi-establishment firms and by providing direct evidence of increased employment at firms' *other* establishments following successful elections. Second, the overall negative effects may be partially driven by managers' or owners' dislike of working with unions, unrelated to unions' costs. Supporting this, we find the largest effects at non-unionized firms and at elections with the longest delay during the election process, both proxies for employers' opposition to unions. This evidence provides new reasons why unionization may lead to establishment closure even without direct wage or productivity effects, which is consistent with past research on these elections. Overall, this shows that any efforts to increase collective bargaining or reform labor law in the U.S. should address employers' ability to avoid unions and their overall opposition to unions.

References

American Compass, The (2021) "Note What they Bargained For: Worker Attitudes About Organized Labor in America."

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.

Arnold, David (2019) "Mergers and Acquisitions, Local Labor Market Concentration, and Worker Outcomes," *Working Paper*.

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*.

Berger, David, Kyle Herkenhoff, and Simon Mongey (2021) "Minimum Wages and Welfare," *Working Papers*, p. 49.

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.

Boehm, Christoph E., Aaron B. Flaaen, and Nitya Pandalai-Nayar (2019) "Input Linkages and the Transmission of Shocks: Firm-Level Evidence from the 2011 Tohoku Earthquake," *Review of Economics and Statistics*, Vol. 101.

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.

Cameron, Andrew Tangel and Doug (2020) "WSJ News Exclusive Boeing to Move All 787 Dreamliner Production to South Carolina," *Wall Street Journal*.

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.

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.

Cohen, Larry and Richard Hurd (1998) "Chapter 11: Fear, Conflict, and Union Organizing," in *Organizing to Win : New Research on Union Strategies*, Ithaca, NY: ILR Press, pp. 181–196.

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.

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, 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.

Edin, Per-Anders and Robert Topel (1997) "Wage Policy and Restructuring: The Swedish Labor Market since 1960," in Richard B. Freeman, Robert H. Topel, and Birgitta Swedenborg eds. *The Welfare State in Transition: Reforming the Swedish Model*, Chicago: University of Chicago Press, pp. 155–201.

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.

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.

Flanagan, Robert (2007) "Has Management Strangled U.S. Unions?" in *What Do Unions Do? A Twenty Year Perspective Edited by James Bennett and Bruce Kaufman*, New York: Routledge.

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.

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.

Friedman, Milton (1951) “Some Comments on the Significance of Labor Unions for Economic Policy,” in *The Impact of the Union*, New York: Harcourt Brace, pp. 204–259.

Ganong, Peter and Pascal Noel (2020) “Liquidity versus Wealth in Household Debt Obligations: Evidence from Housing Policy in the Great Recession,” *American Economic Review*, Vol. 110, pp. 3100–3138.

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*.

Gardner, John (2021) “Two-stage differences in differences,” *Working Paper*.

Garwin Corp, 153 N.L.R.B 1030 (1968) .

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.

Goncalves, Felipe (2021) “Do Police Unions Increase Misconduct?” *Working Paper*, p. 76.

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.

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.

Harju, Jarkko, Simon Jäger, and Benjamin Schoefer (2021) “Voice at Work,” *Working Paper*, p. 62.

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

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*.

Jensen, Michael and William Meckling (1976) “Theory Of The Firm: Managerial Behavior, Agency Costs And Ownership Structure,” *Journal of Financial Economics*, Vol. 3, pp. 305–360.

Kahn-Lang, Ariella and Kevin Lang (2020) “The Promise and Pitfalls of Differences-in-Differences: Reflections on *16 and Pregnant* and Other Applications,” *Journal of Business & Economic Statistics*, Vol. 38, pp. 613–620.

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.

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 (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.

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.

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 (2021) “When Is Parallel Trends Sensitive to Functional Form?” *arXiv:2010.04814 [econ, stat]*.

Russ, Hilary (2022) “Starbucks union claims company closed two cafes in retaliation,” *Reuters*.

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.

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.

Slichter, Sumner, James Healey, and Robert Livernash (1960) *The Impact of Collective Bargaining on Management*, Washington DC: Brookings Institution.

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.

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

——— (1955b) "Marshall and Friedman on Union Strength," *The Review of Economics and Statistics*, Vol. 37, p. 384.

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.

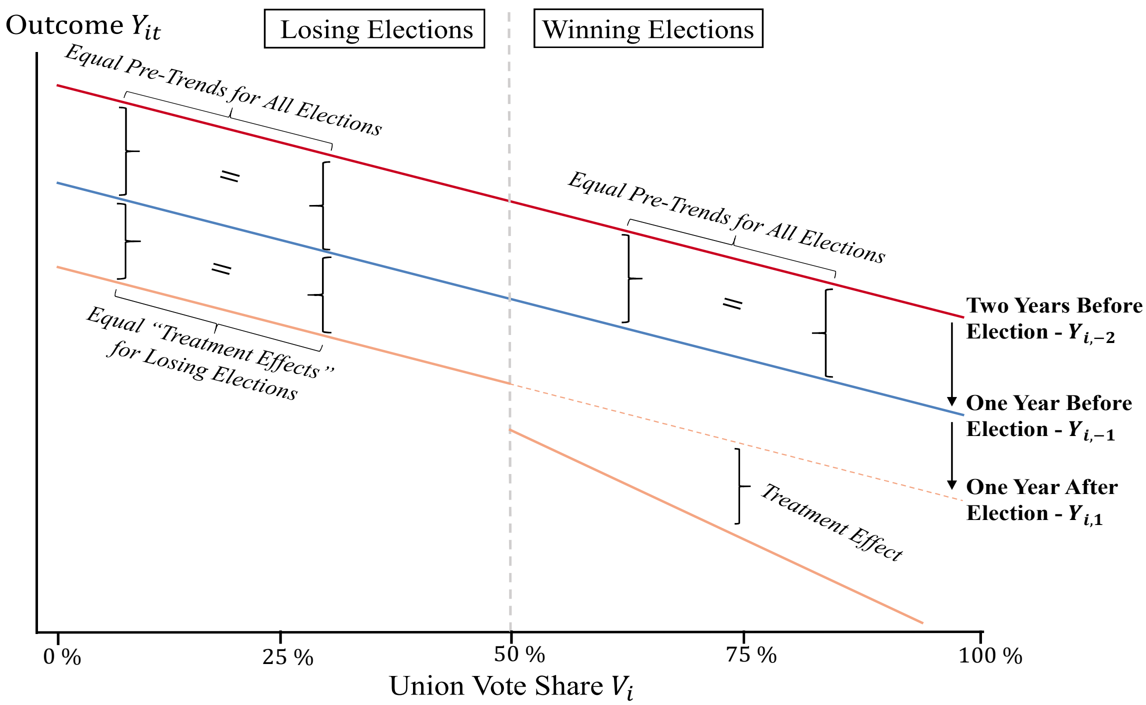
Weiler, Paul (1983) "Promises to Keep: Securing Workers' Rights to Self-Organization under the NLRA," *Harvard Law Review*, Vol. 96, p. 1769.

Wright Line, 251 N.L.R.B 1083 (1980) .

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

9 Figures

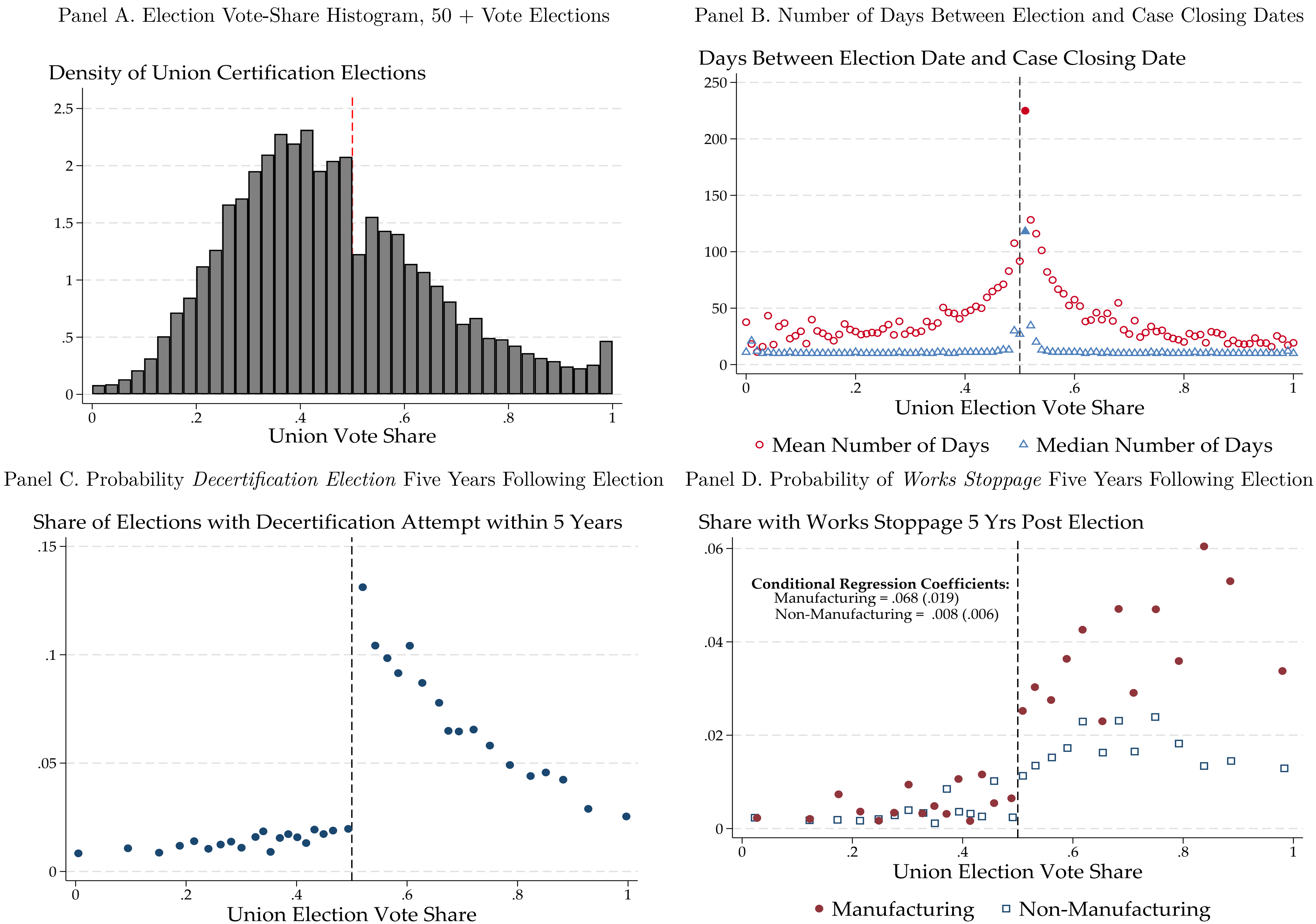
Figure 1: Testable Implications of Parallel Trends Identifying Assumption



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

Figure 2 Note: Figure 2 presents four panels illustrating characteristics of close union elections. All panels are constructed using external union election data (e.g., not our final sample matched to the Census) but the sample was constructed to mirror the overall sample construction (see Appendix 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. Panel C plots the probability of each union election experiences a decertification election in the five years following the case closing. The decertification elections are also from our combined NRLB datasets but excluded from our main analysis. Panel D plots the probability of each union election experiencing a works stoppage in the five years following the case closing. The works stoppage data is from the FMCS and covers works stoppages from 1984–2019. Consequently, we only plot follow-up works stoppages for elections from 1984–2005. For the decertification and works stoppage figures, we match based on exact company names and cities rather than the SoftTFIDF algorithm we use for the main analysis. The "conditional regression coefficients" are the coefficients from regressing the stoppage indicators on the vote share for winning elections including controls for deciles of the number of workers in the bargaining unit, the four-digit NAICS industry, and election state.

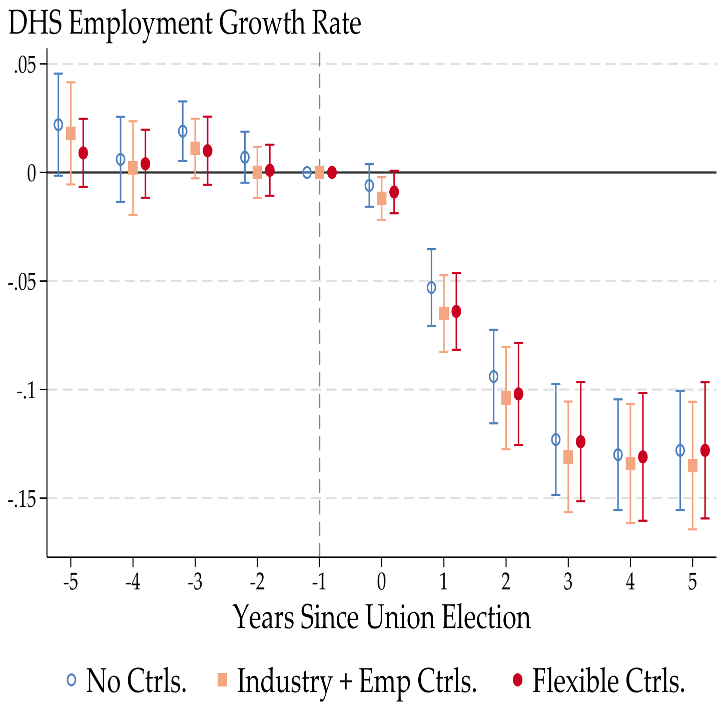
Figure 2: Characteristics of Close Elections that Motivate Including Larger Margin-of-Support Elections



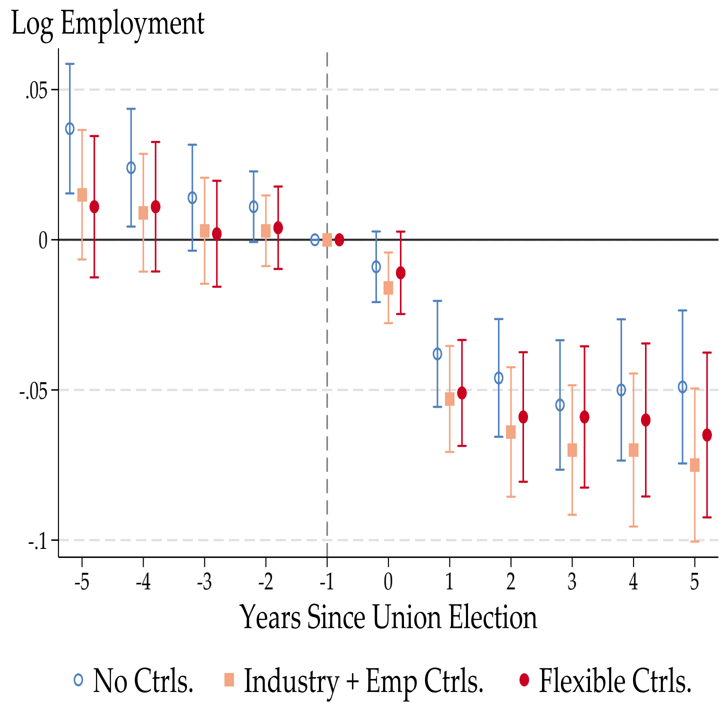
Note: See the previous page.

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

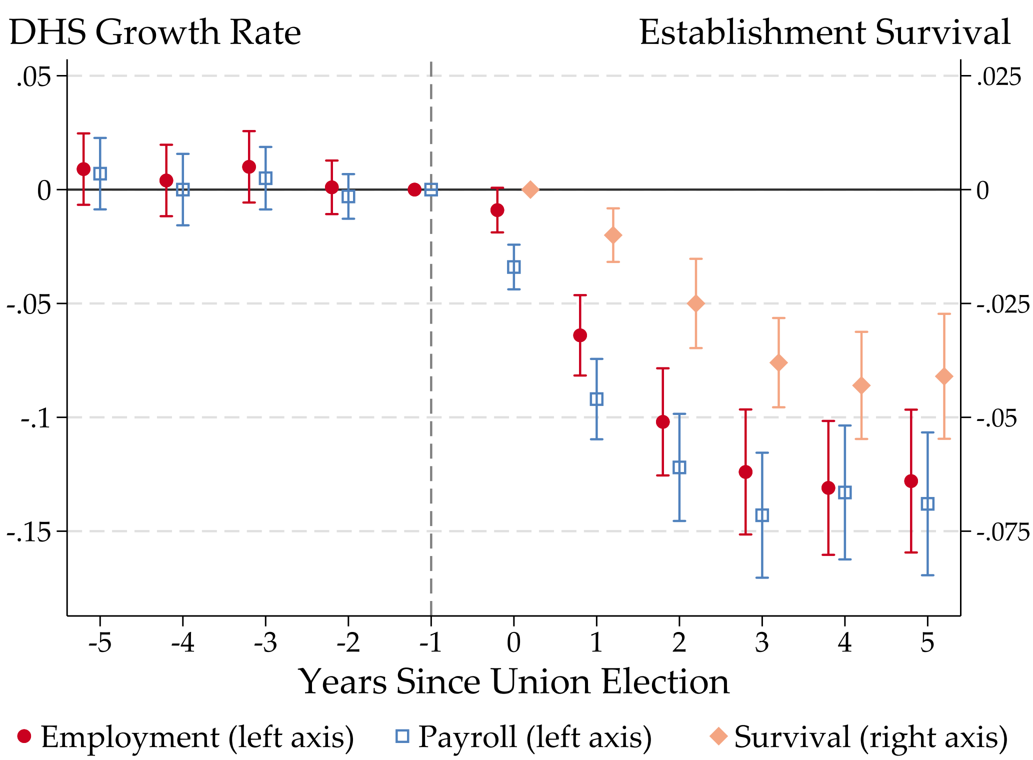
Panel A. DHS Employment Growth



Panel B. Log Employment

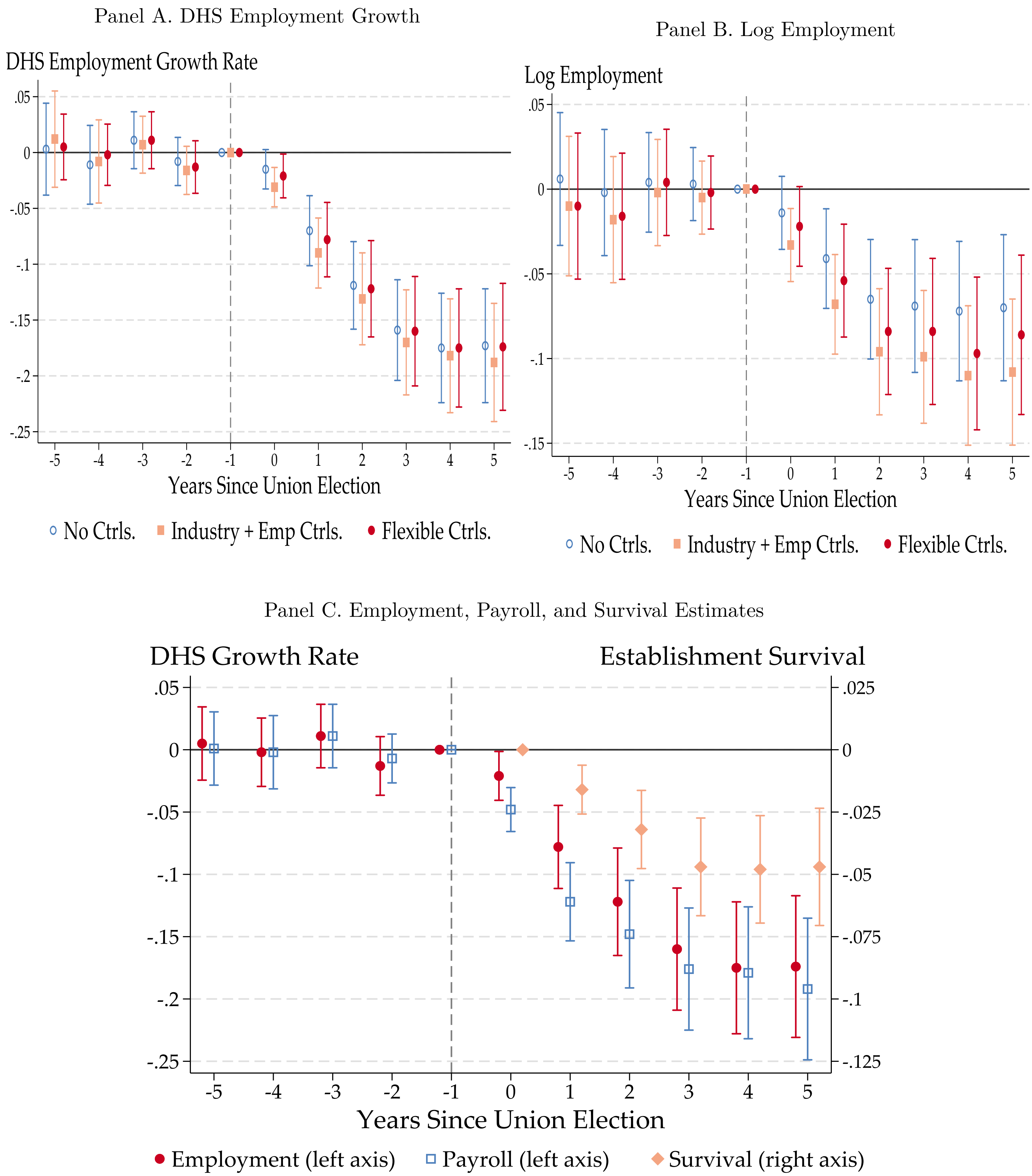


Panel C. Employment, Payroll, and Survival Estimates



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 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).

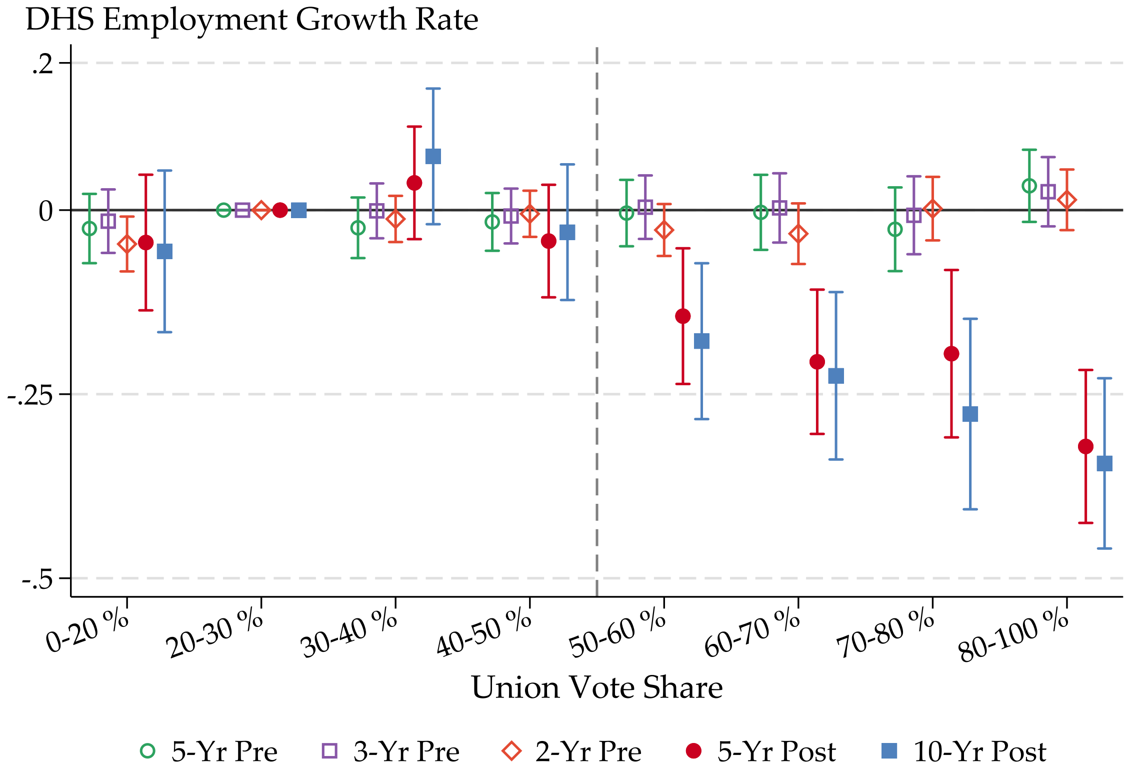
Figure 4: Employment and Survival Estimates, 20–80% Vote-Share Elections, Manufacturing



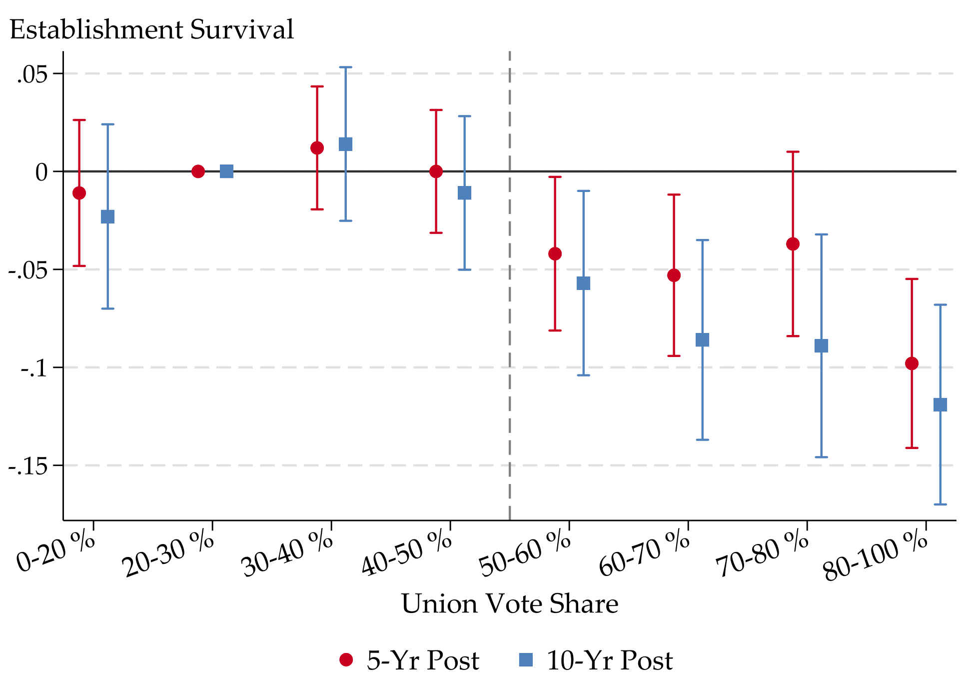
Note: These estimates are identical to Figure 3 except that they are only estimated for manufacturing elections.

Figure 5: Nonparametric Vote-Share Heterogeneity Estimates, Manufacturing

Panel A. DHS Employment Growth Rate



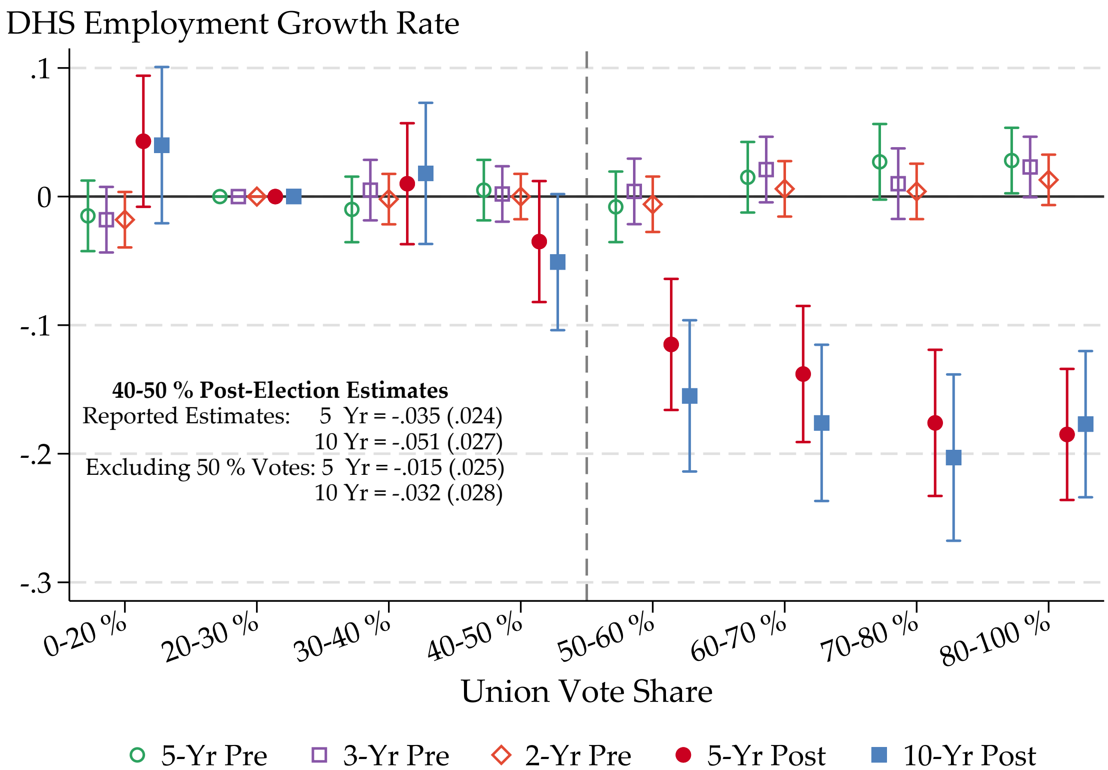
Panel B. Establishment Survival



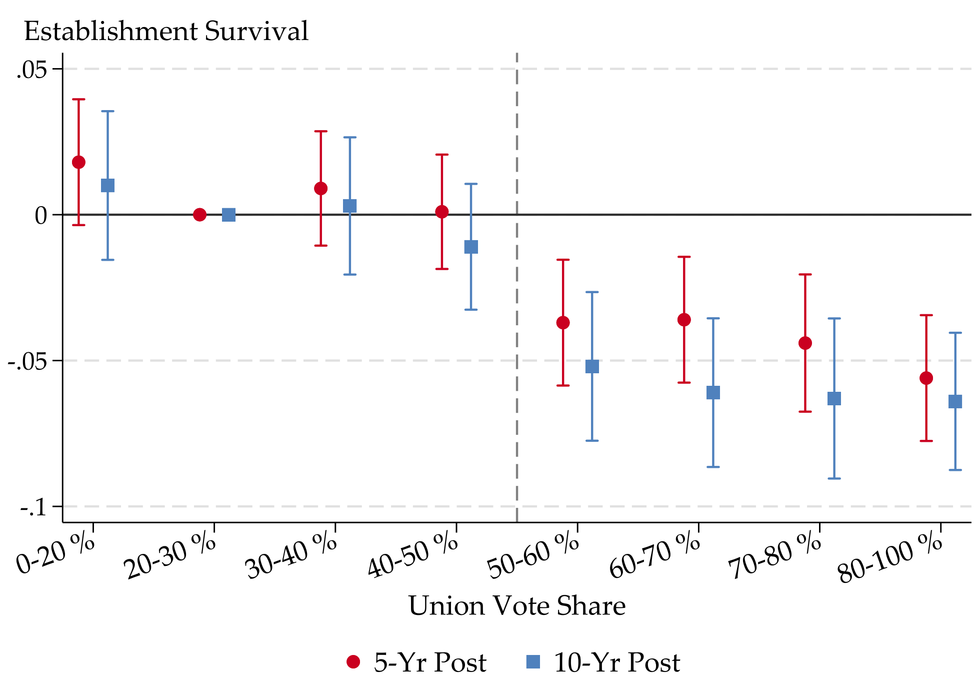
Note: This figure plots the $\delta_{g,n}$ coefficients from estimating the vote-share heterogeneity specification 10 with the vote-share distribution partitioned into eight groups indicated on the x-axis. We omit the 20–30% election group so the other estimates are relative to that group. The sample 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 for establishment survival. The estimates include the flexible control specification (see Section 4 for details). Standard errors are clustered by establishments' firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment).

Figure 6: Nonparametric Vote-Share Heterogeneity Estimates, All Industries

Panel A. DHS Employment Growth Rate

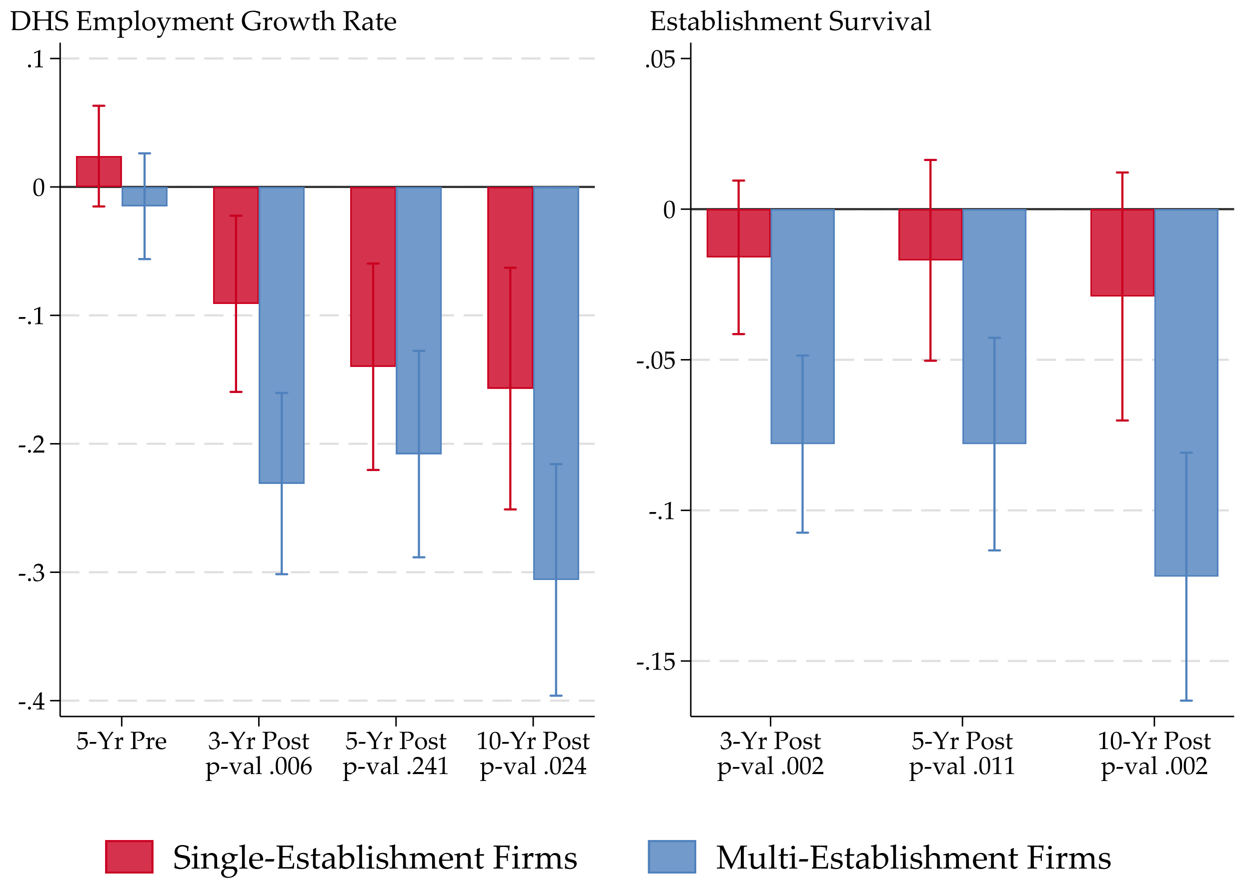


Panel B. Establishment Survival



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).

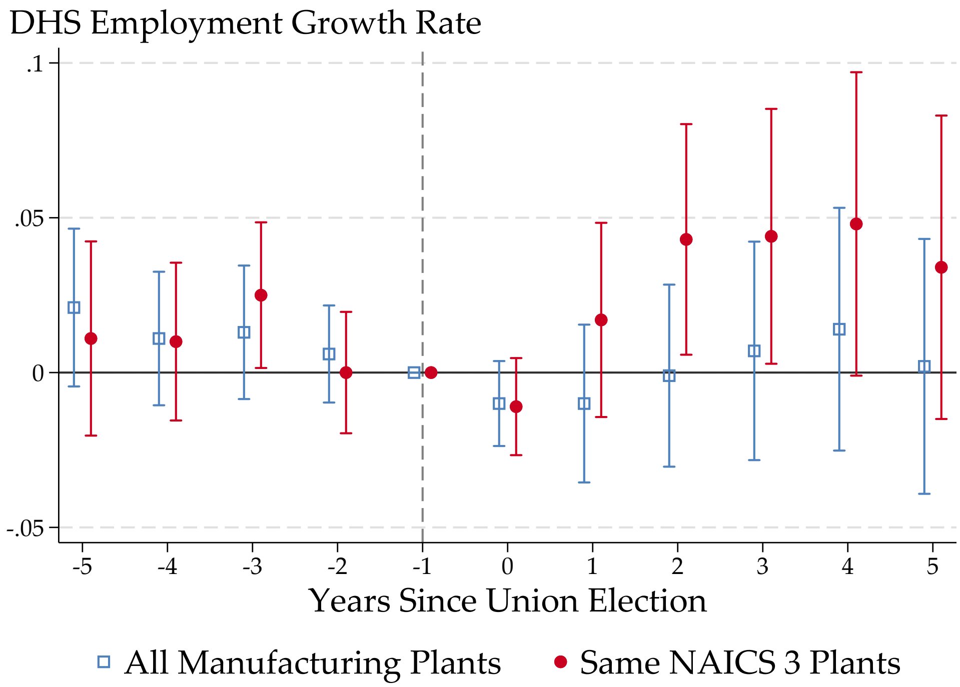
Figure 7: Single- Versus Multi-Establishment Firm Heterogeneity



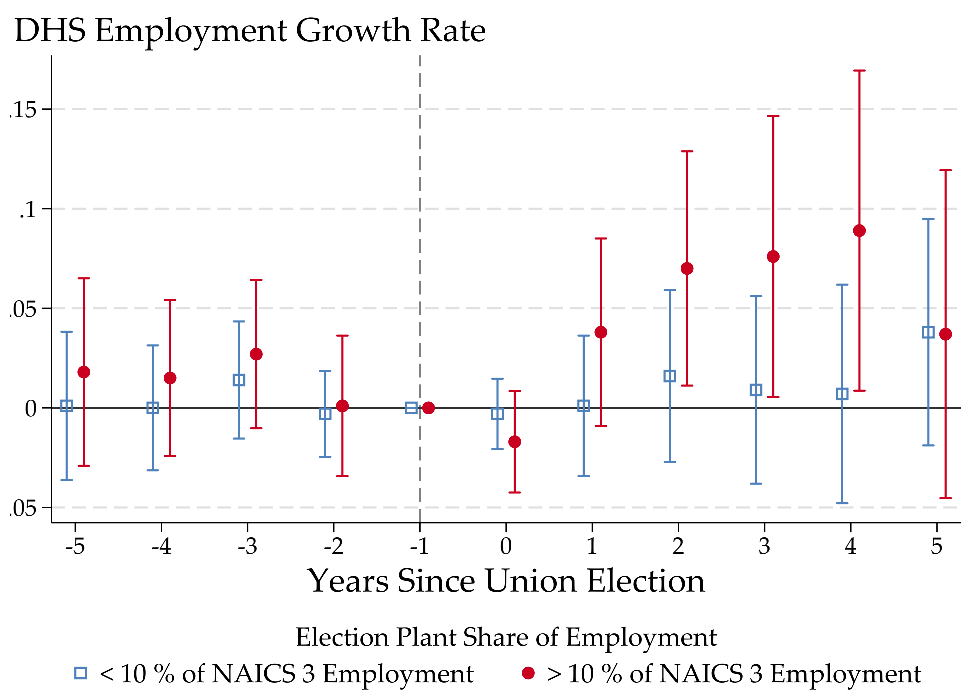
Note: This figure plots the $\delta_{h,n}$ coefficients from estimating our heterogeneity specification in equation 12 for elections at single- versus multi-establishment firms. An election at a multi-establishment firm is defined based on whether the establishment's firm has any other establishments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

Figure 8: Employment Effects of Successful Elections on Firms' Other Establishments

Panel A. All Establishments and Three-Digit NAICS Establishments

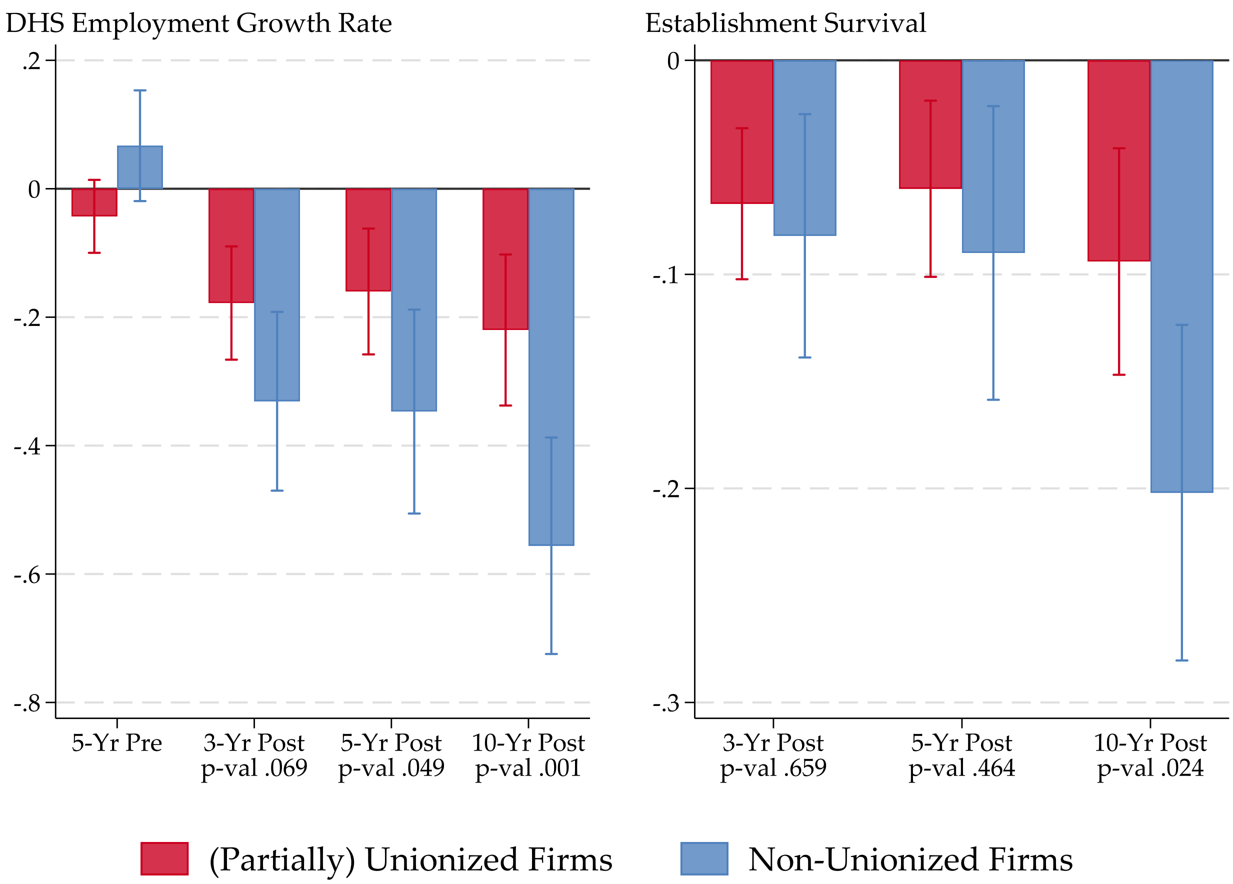


Panel B. Estimates by Election's Employment Share



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

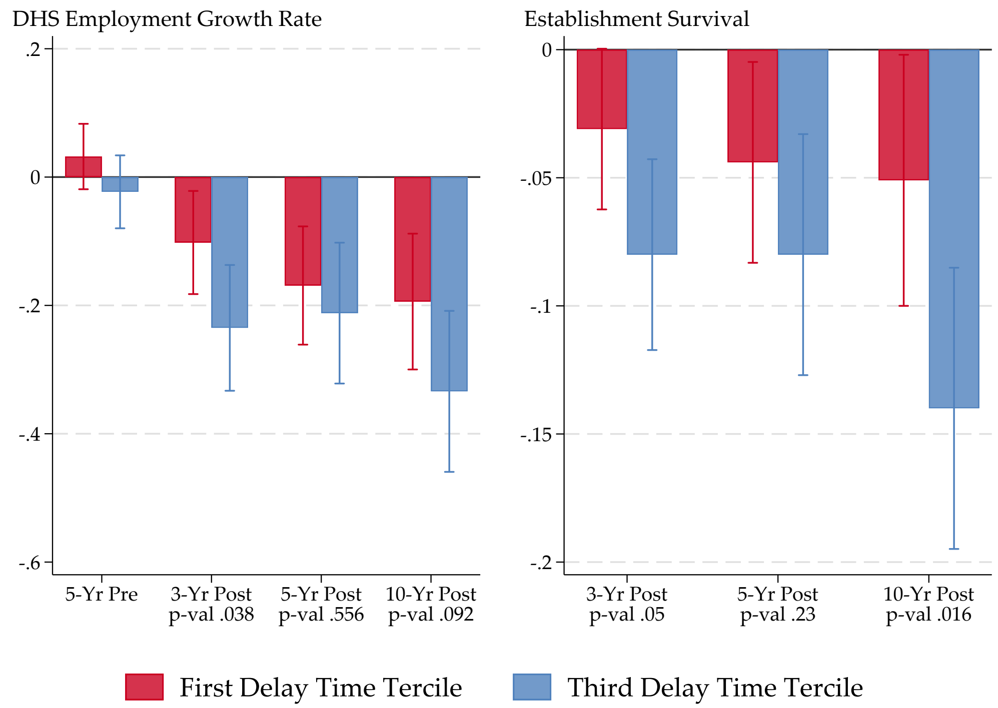
Figure 9: Unionized versus Non-Unionized Firm Heterogeneity



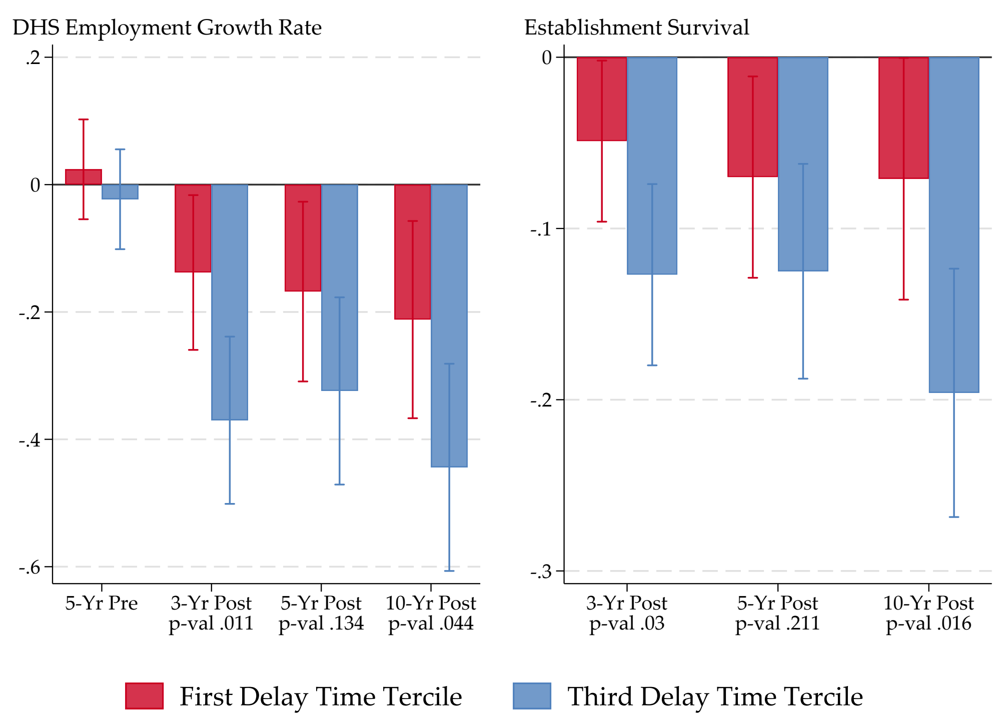
Note: This figure plots similar heterogeneity results as Figure 7 except that the heterogeneity is for elections at multi-establishment firms with at least one unionized establishment versus firms without any unionized establishments. See Appendix D for how we define firms' unionization status. The controls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

Figure 10: Election Delay Heterogeneity

Panel A. All Elections

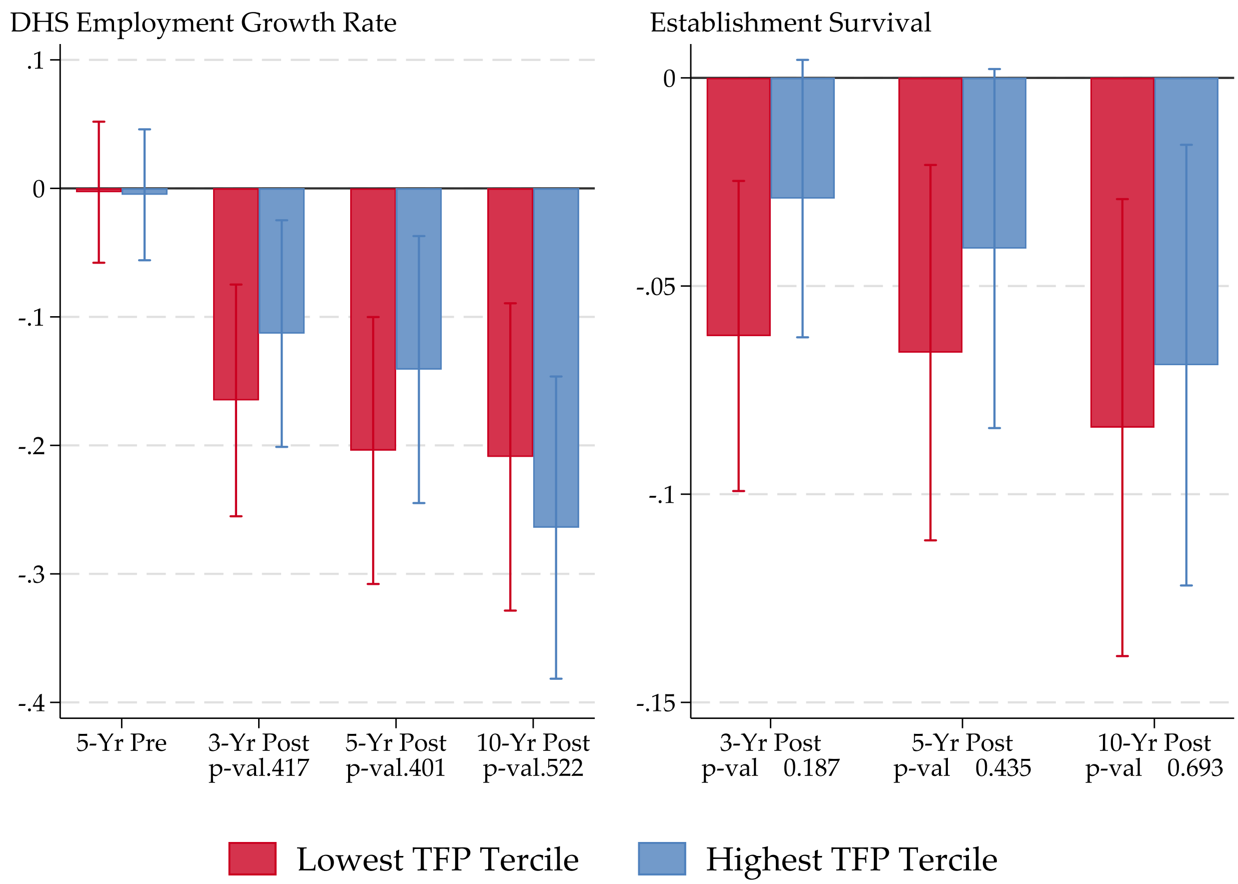


Panel B. Elections at Multi-Establishment Firms



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

Figure 11: Establishment-Level Total Factor Productivity Heterogeneity



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

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 round to the nearest 1,000.

Table 2: 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 11). A five-year coefficient of 0.03 implies that elections with 75% of votes grew approximately 1.5 percent slower during the five years before the election than an election with 25% of votes. The outcome for all specifications is establishment-level DHS employment growth relative to time -1. The sample includes 20–80% vote-share elections. The first two columns include elections in all industries and the last two columns include just manufacturing elections. The odd columns include only industry and employment controls and the even columns include our flexible control specification (see Section 4 for details). Standard errors are clustered by establishments' firmid during the year of the election (e.g., the clustering variable is fixed over time for each establishment). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: **Post-Election Outcome Trends by Vote Share, 20–80% Vote-Share Elections**

Outcome:	DHS Emp Growth Rate		Establishment Survival	
Panel A: All Industries				
Event-Time × 0-50 % Vote Share				
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				
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				
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				
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 11 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 A2 for the same results with alternative included controls. The columns that "Exclude 50% Elections" include an interaction between having a vote share of exactly 50% and event time. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Employment and Survival Estimates by Industry, 20–80% Vote-Share Elections

Industry Group:	Manufacturing		Services		Other	
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	6,000	6,000	8,000	8,000	5,000	5,000
Industry Group Share of Elections	0.302	0.302	0.414	0.414	0.284	0.284

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

Table 5: 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 + Employment Ctrls.	X	X
Flexible Ctrls.	X	X

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

Table 6: Election Delay Heterogeneity, Continuous Delay Time Specification

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**	-0.045**	-0.120**	-0.055**
	(0.058)	(0.023)	(0.055)	(0.022)	(0.059)	(0.023)
5-Year Post Election	-0.121*	-0.064**	-0.104*	-0.051**	-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**	-0.074**	-0.132*	-0.061*
	(0.073)	(0.033)	(0.073)	(0.032)	(0.074)	(0.033)
Industry + Employment Ctrls.	X	X	X	X	X	X
Pooled Ctrls.			X	X		
Flexible Ctrls.	X	X			X	X
Observations	6,000	6,000	6,000	6,000	6,000	6,000

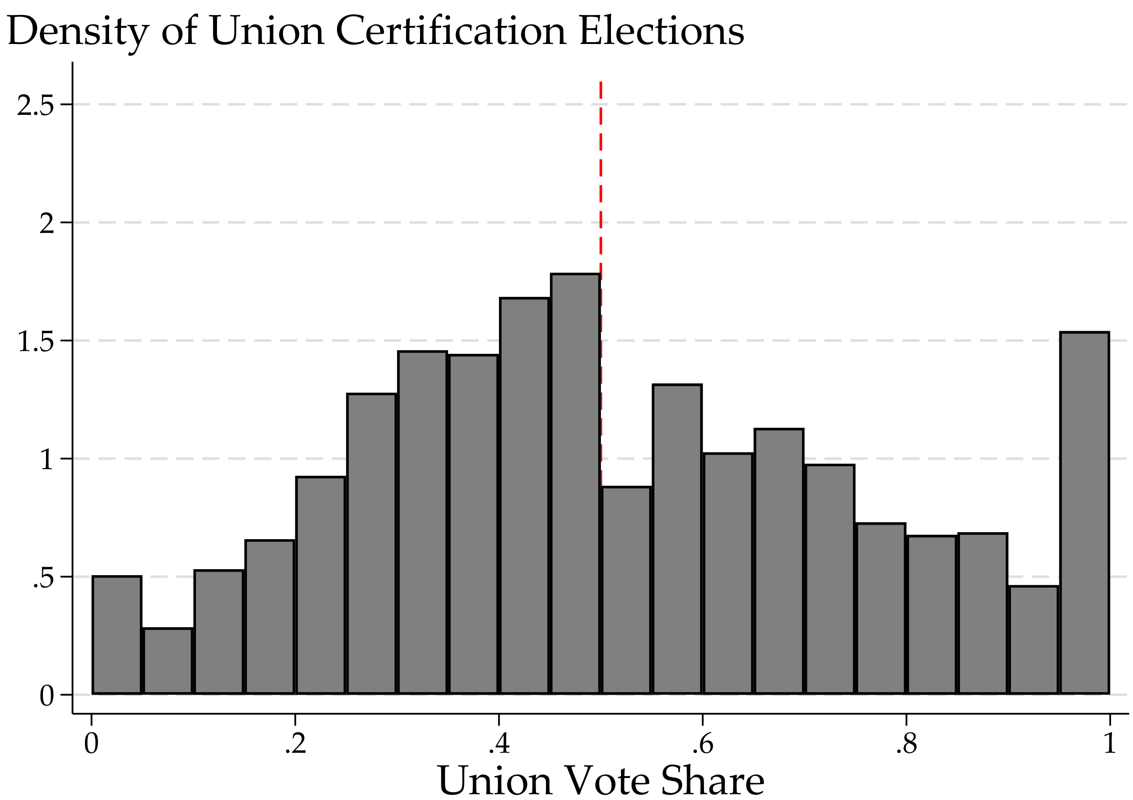
Note: This table presents coefficient estimates from a modified version of specification 8. Specifically, we interact the treatment by event time indicators with the continuous *log delay time*. See Appendix D for details on how we calculate the delay time. The table reports the coefficients on these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

Appendix for

Unionization, Employer Opposition, and Establishment Closure

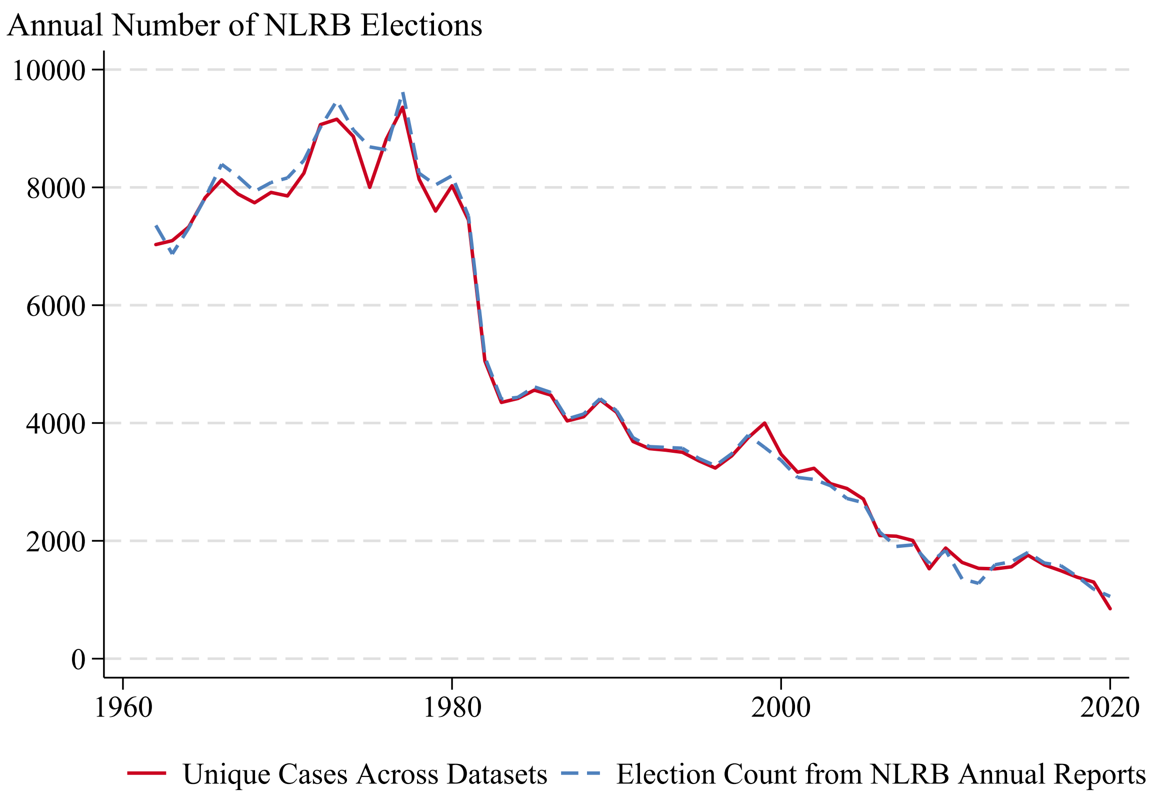
A Appendix Figures

Figure A1: 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 2 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.

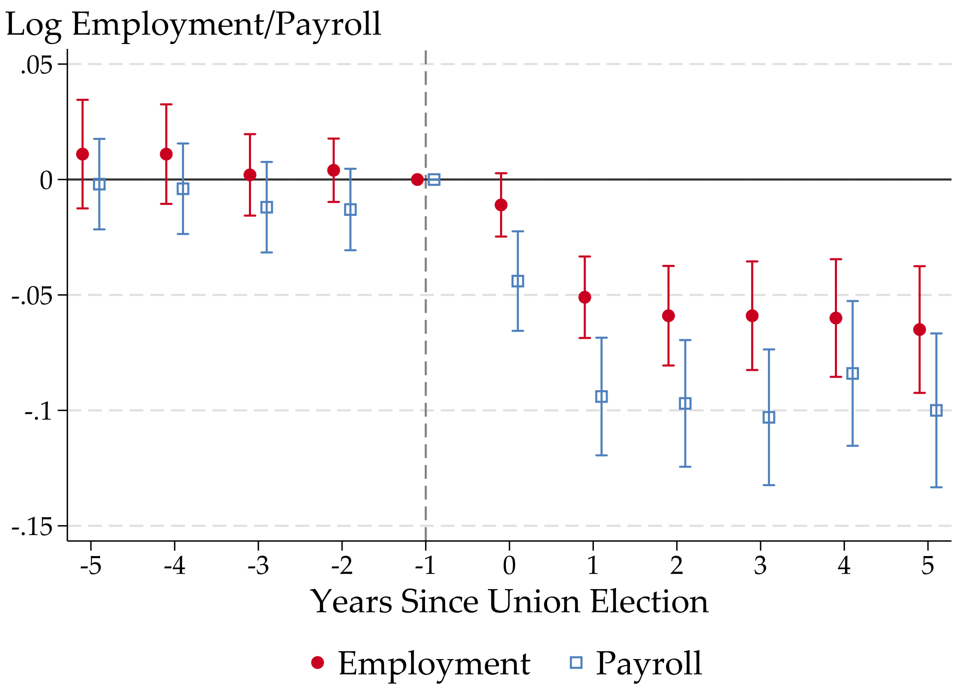
Figure A2: Number of Unique Case Numbers Across Datasets versus NLRB Annual Reports



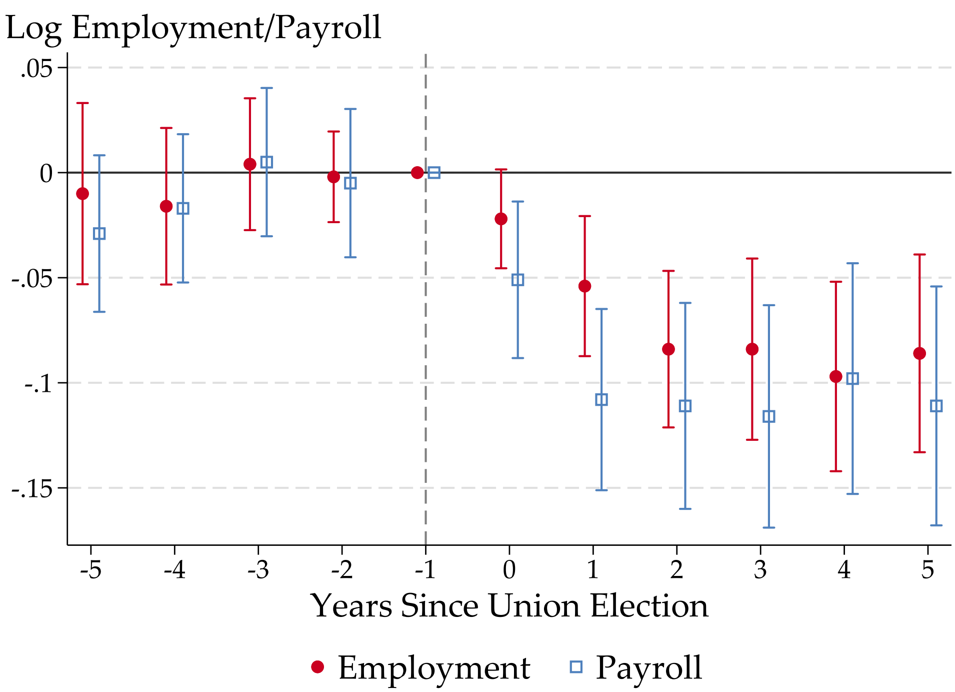
Note: This figure plots the total number of unique NLRB election cases each year in our dataset and in the annual NLRB reports. These include all case types (e.g., 'RC' cases and non-RC cases) Our dataset is from combining union election datasets from Henry Farber, J.P. Ferguson, and Thomas Holmes and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

Figure A3: Log Employment and Payroll Estimates, 20–80% Vote-Share Elections

Panel A. All Union Elections

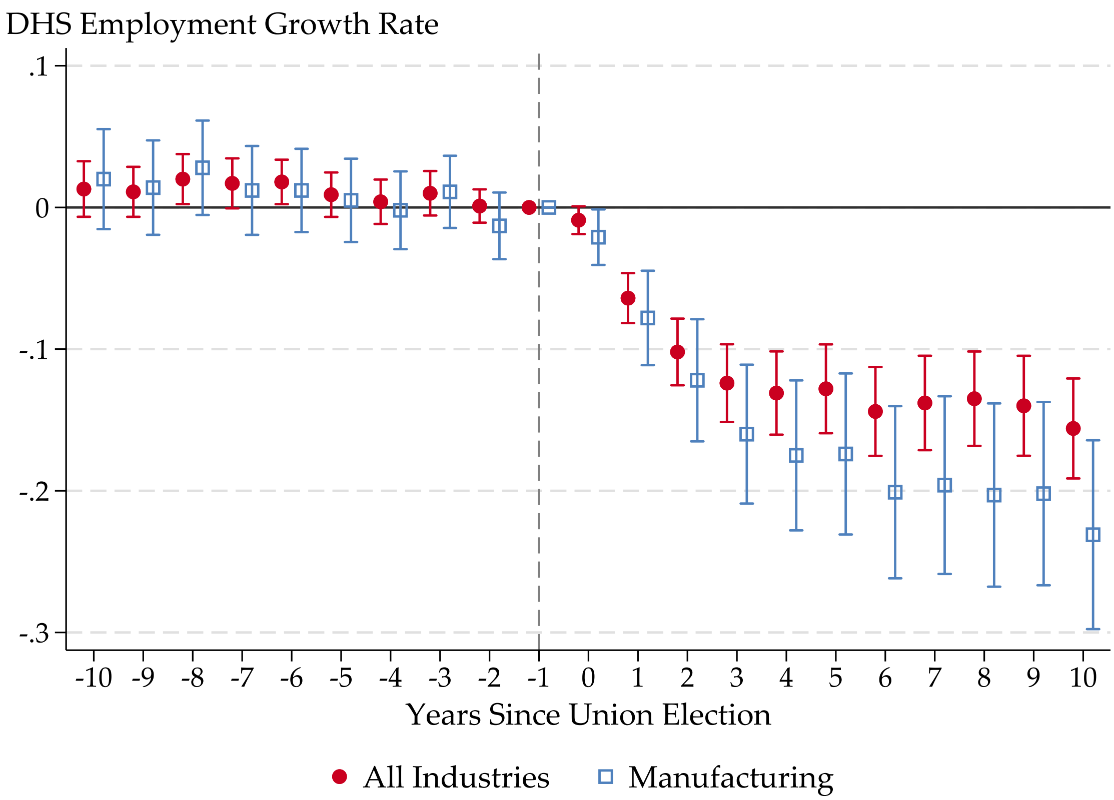


Panel B. Manufacturing Union Elections



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.

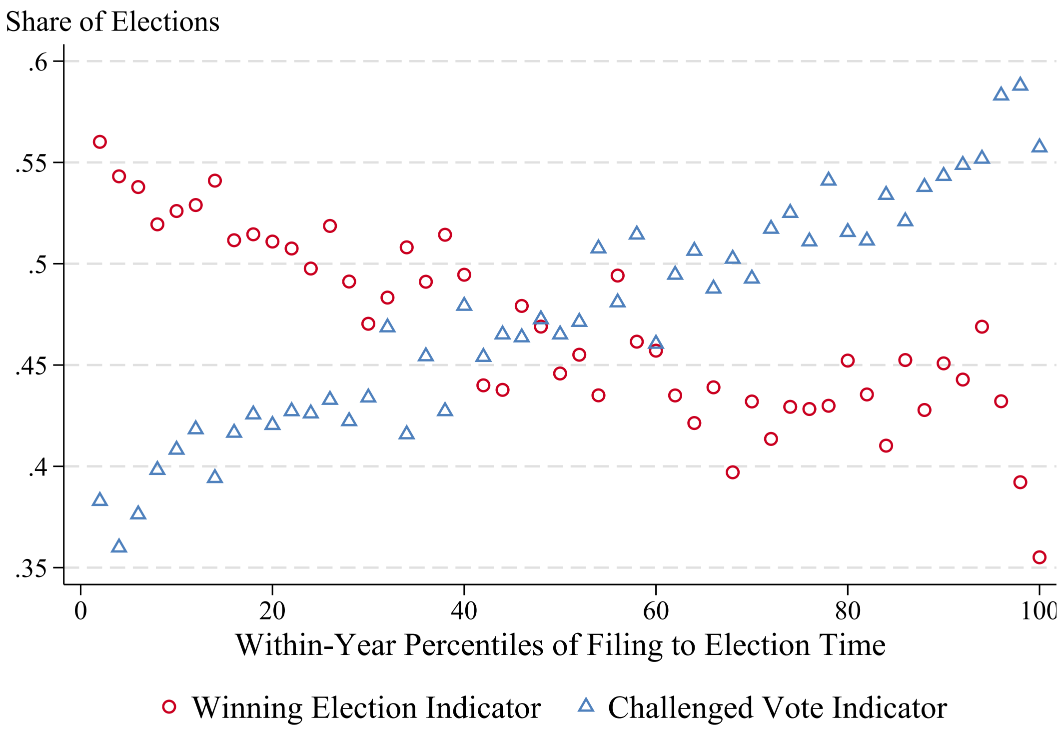
Figure A4: DHS Employment Estimates, 20–80% Vote-Share Elections, 10 Yr Pre- and Post-Periods



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 but not from -10 to -5 years pre-election. Consequently, each of the -5 to -10 point estimates average over slightly different cohorts.

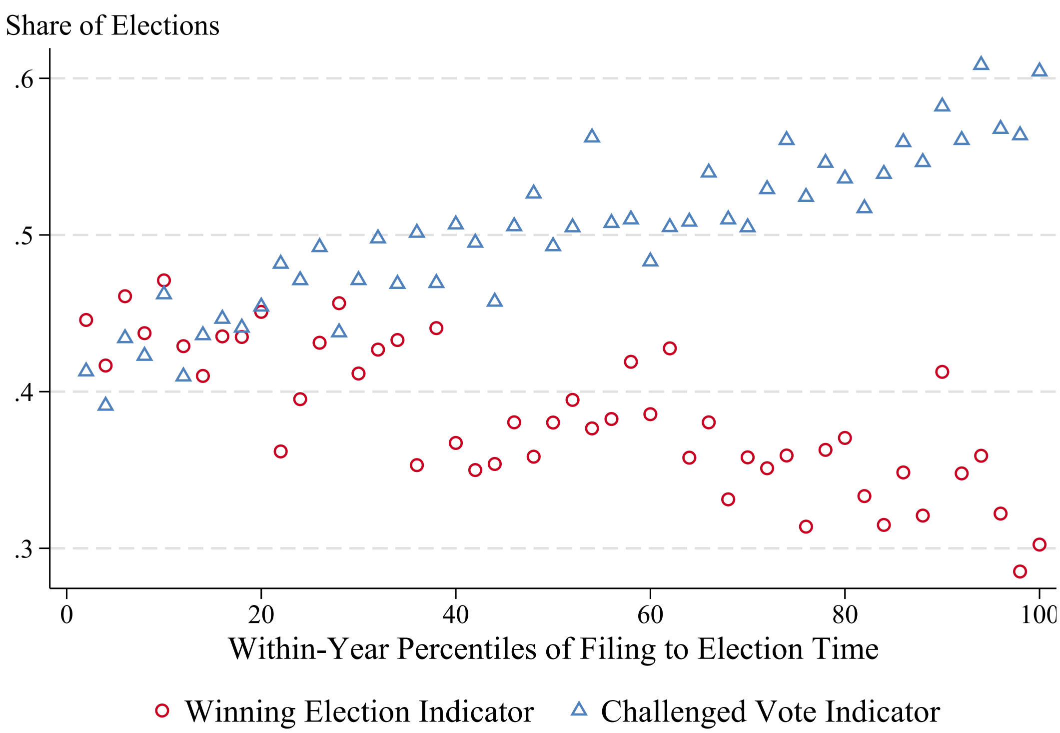
Figure A5: Election Win Rates and Challenged Vote Rates by Delay Time

Panel A. All Elections



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

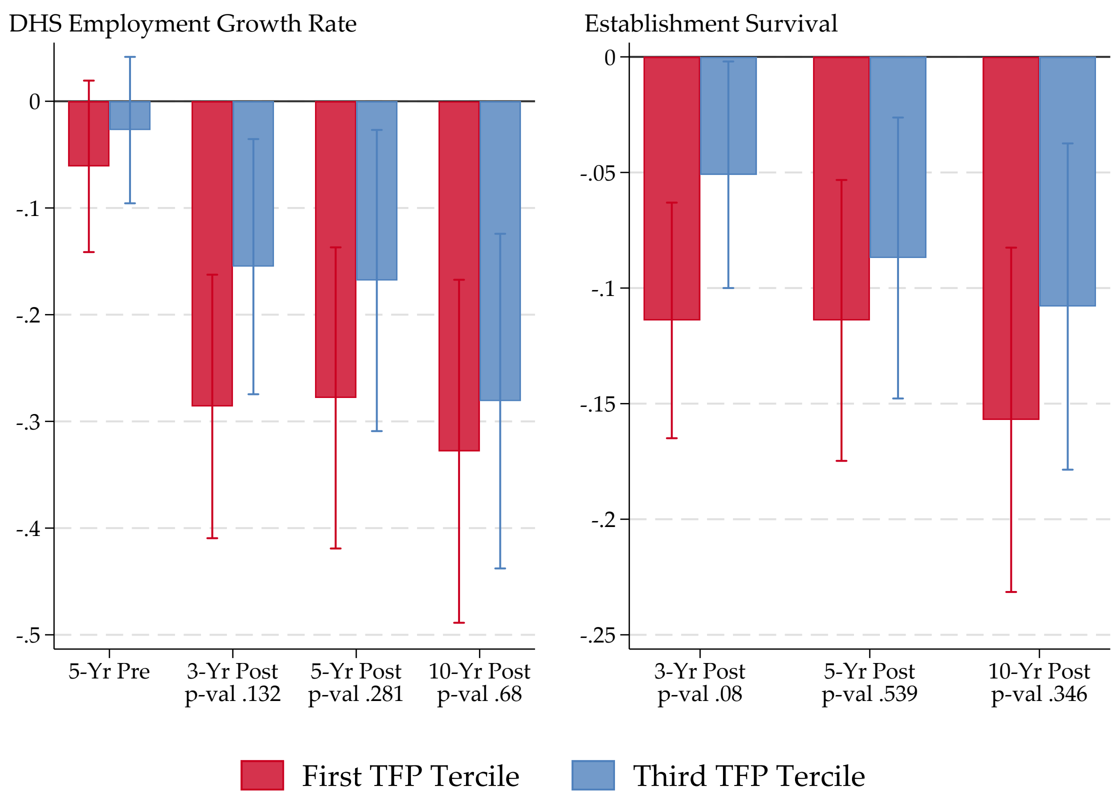
Panel B. Manufacturing



Conditional Regression Coefficients x 100: Winning = -.063 (.013). Challenged = .054 (.013).

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 A6: Establishment-Level Total Factor Productivity Heterogeneity, Multi-Establishment Firms



Note: This figure plots the same estimates as in Figure 11 except restricting the TFP comparison to only be between multi-establishment firms in different TFP terciles. As with the other heterogeneity tests, the sample includes all manufacturing elections and pools the controls across the entire sample.

B Appendix Tables

Table A1: Union Election Matched Sample Construction

	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 without more than five eligible voters. The fourth row only includes non-contested elections (e.g., elections with one union on the ballot). Panel B presents the industry composition of the remaining elections from the fourth row of Panel A. Note we use the NLRB election industry codes here rather than the LBD industry codes but the overall industry shares are reassuringly similar to the industry shares in Table 4. The three columns represent the total shares of elections and eligible voters for all elections and winning elections. Panel C shows our final sample sizes from the matched Census data. The sample restrictions between "Elections Matched to Census Establishments" and "Final Establishment-LEvel Outcome Sample" include keeping (1) the first election at each establishment, (2) at least three years of pre-election survival, (3) non-missing employment, payroll, and other controls at event time $t = -1$.

Table A2: 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	-0.134	0.021	-0.181	0.042
	(0.100)	(0.037)	(0.159)	(0.059)
Event-Time × 50-100 % Vote Share	-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	-0.119	-0.009	-0.150	0.023
	(0.116)	(0.047)	(0.187)	(0.076)
Event-Time × 50-100 % Vote Share	-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	-0.218	-0.052	-0.186	-0.020
	(0.133)	(0.057)	(0.224)	(0.097)
Event-Time × 50-100 % Vote Share	-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 Tables 3 but only includes the baseline industry and employment controls. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Employment and Survival Bargaining Unit Share Interaction, 20–80% Vote-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, 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 A4: Manufacturing versus Services Employment and Survival Estimates, Robustness Checks

Specification:	Baseline		Pooled Controls		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	-0.118***	-0.022	-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***	-0.159***	-0.05***	-0.196***	-0.061***	-0.193***	-0.06***	-0.171***	-0.054**
	(0.043)	(0.019)	(0.04)	(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
Pooled Ctrls.			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 4. Specifically, it presents the differences between the five- and ten-year DHS employment growth rate and survival estimates for various alternative specifications. The first two columns present the differences for the estimates presented in Table 4. The "Pooled Controls" columns pool the controls across all cohorts as described in Section 4. The "Good Matches" columns restrict to election matches which we give a 95% rating (see Appendix D for details). The "Barg Unit Share" columns restrict to elections where the bargaining unit is at least 25% of the total establishment employment. The 30–70% columns restrict to elections with 30–70% of the vote share. For all specifications with restrictions, we still use the entire sample for controls but restrict the treated variables to be estimated from the restricted sample. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

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

Specification:	Baseline		Pooled Controls		30-70 %	
Outcome:	DHS Emp	Survival	DHS Emp	Survival	DHS Emp	Survival
5-Year Difference	-0.068	-0.061**	-0.063	-0.053**	-0.034	-0.057**
	(0.058)	(0.024)	(0.054)	(0.023)	(0.065)	(0.027)
10-Year Difference	-0.149**	-0.093***	-0.13**	-0.085***	-0.067	-0.06*
	(0.066)	(0.03)	(0.062)	(0.028)	(0.075)	(0.034)
Industry + Employment Ctrls.	X	X	X	X	X	X
Pooled Ctrls.			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 7. Specifically, it presents the differences between the five- and ten-year DHS employment growth rate and survival estimates for various alternative specifications. The first two columns present the differences for the estimates presented in Figure 7. The "Pooled Controls" columns pool the controls across all cohorts as described in Section 4. The 30–70% columns restrict to elections with 30–70% of the vote share. For all specifications with restrictions, we still use the entire sample to estimate controls but restrict the treated variables to be estimated from the restricted sample. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Unionized versus Non-Unionized Firm Heterogeneity, Robustness Checks

Specification:	Baseline		Pooled Controls		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	-0.139	-0.018	-0.179	-0.057	-0.197*	-0.029
	(0.095)	(0.041)	(0.089)	(0.039)	(0.112)	(0.049)	(0.108)	(0.046)
10-Year Difference	-0.336***	-0.108**	-0.287***	-0.09**	-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
Pooled Ctrls.			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 9. Specifically, it presents the differences between the five- and ten-year DHS employment growth rate and survival estimates for various alternative specifications. The first two columns present the differences for the estimates presented in Figure 9. The "Pooled Controls" columns pool the controls across all cohorts as described in Section 4. The "Contracts since 1990" column only classifies firms as unionized versus non-unionized starting in 1990. This gives all firms at least five years of pre-election FMCS contract data that we can use to define the firms' unionization status. The 30–70% columns restrict to elections with 30–70% of the vote share. For all specifications with restrictions, we still use the entire sample to estimate controls but restrict the treated variables to be estimated from the restricted sample. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

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.[57] 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

[57] 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.

not extrapolate to elections that win by larger margins of victory. However, we find large survival 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 for 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 wages 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 much 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). Second, 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 (e.g., more productive workers become union members). 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).[58]

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.

[58]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.

(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 57 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 wages 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 because the 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. 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.[59] Finally, Galdon-Sanchez and Schmitz (2002) and Schmitz (2005) document how

[59] According to the FMCS works stoppage data, the Bridgestone and Catepillar 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

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.[60] Additionally, there might be separate elections for multiple different bargaining units filed under the same case number (e.g., if a union initially filed for a petition for one bargaining unit but the NLRB then split 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 somewhat 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 there are duplicate observations within the same case number. For these datasets, we first pick the observation with the last election date and then the observation with the largest bargaining unit size.

This leaves us one observation per case number within each dataset but duplicates across datasets. We take one observation per case number across datasets. For picking a single case number

strikes following successful union elections in our sample displayed in Figure 2, the median (mean) duration of works 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.

[60] 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

per dataset, we deprioritize observations in the Farber data given data 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 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 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.[61] This suggests that the employers' 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 address information in the bargaining unit description and consequently use the election site addresses when they 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[62]. 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%).

- **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

[61] 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 employers' location is different than the employees' worksite (e.g., security guards), the election might be held at the work site

[62] 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

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 2 and A5 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 % 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 be we always only match each contract to one establishment.[63]

There are duplicate observations both across the Holmes versus FMCS datasets and within each dataset.[64] 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.

FMCS Works' Stoppage Data

For Figure 2 Panel D, we use works' stoppage data from the FMCS from 1984–2005. The data are available here. They include both strikes and employer-initiated lockouts. We match the works 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 works' stoppage data.

[63]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.

[64]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.

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 941s* that provide annual employment and payroll at the EIN-level which can cover multiple establishments. The Census uses an imputation model to allocate these EIN-level measures across establishments. This model primarily imputes employment changes across establishments based on their past employment. Consequently, employment changes at an establishment 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 Manufacturers (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 Manufacturers

We define establishment-level productivity using inputs and outputs from the Annual Survey of Manufacturers (ASM) and TFP measures calculated by Foster et al. (2016). To classify each election into different terciles of the 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 its $E_i - 2$ productivity rank). Based on the election observations with defined TFP, we then classify the elections into within-year terciles based on these rankings.

Matching Elections, Contracts, and LBD 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 use to directly link the establishments to administrative Census data firms. 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 compnay names (e.g., `MANUFACTURERS` ⇒ `MANUFACTURING`).

Election, Contract, and Census Address Geocodes: We geocode all 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 similarly 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 an observations' street address is missing or we cannot geocode it, we take the city/state geocode or the zip-code geocode.

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 instead 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.[65]

2. For each election-establishment pair, we calculate the Soft TF-IDF similarly 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

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

[65] 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

where weight(w, A_j) is defined as

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

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

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

Intuitively, the `TF` portion of the weight gives higher weights to words 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 geocoordinates 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 similarly 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.[66] 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 records where the name and geographic location match exactly are "true" matches, which correspondingly allows us to estimate that a pair of records with a match score of s is matched correctly with probability:

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

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.

[66]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.

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.

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.

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.

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.

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.