

# Do Unions Cause Business Failures?\*

John DiNardo

University of Michigan, Ann Arbor  
and NBER

David S. Lee

UC Berkeley  
and NBER

March 2003

## Abstract

Estimating the causal effect of unionization on business survival rates is difficult in the absence of large, representative data on establishments with union status information. It is also confounded by selection bias, because unions may tend to organize at highly profitable enterprises that are more likely to survive. Using new data on more than 27,000 establishments that faced organizing drives in the U.S. during 1983-1999, this paper utilizes a regression discontinuity design to estimate the impact of unionization on the probability of employer dislocation. Survival probabilities of employers where unions *barely* won the election (e.g. by one vote) are compared to those where the unions *barely* lost. The analysis yields a surprising result: *little or no* union effect on business dislocation rates over 1- to 18-year horizons.

---

\* An earlier version of the paper "The Impact of Unionization on Establishment Closure: A Regression Discontinuity Analysis of Representation Elections" is on-line as *NBER Working Paper #8993*, June 2002. We thank David Card, Larry Katz, Enrico Moretti, Morris Kleiner, participants of the University of Michigan Labor Workshop and the NBER Labor Studies Summer Institute for helpful comments and suggestions, Matthew Butler and Francisco Martorell for outstanding research assistance, Hank Farber for providing election data, and Christina Lee for reading previous drafts.

# 1 Introduction

It is widely understood that unions raise the cost of labor by raising members' wages above market rates.<sup>1</sup> Unions also impose other costs on employers - limiting discretion in hiring and firing, for example, and altering the structure of pay differentials across skill groups. A key question for understanding the social costs of unionization is whether the wage premiums and other costs of unionism create large or small distortions in the allocation of labor.<sup>2</sup> These distortions can take the form of reduced employment at unionized firms, or most dramatically, an accelerated pace of business failures.

The potentially adverse effects of unions on firm survival are acknowledged by employers and employees alike. During union organizing drives, firms routinely threaten to close a plant if the union drive is successful [Bronfenbrenner 2000]. Employees seem to take these threats seriously: the risk of plant closure is cited as the leading cause of union withdrawal from organizing attempts [Commission for Labor Cooperation 1997]. Such risks are arguably higher now, in light of rapidly expanding trade with low-wage countries such as China and Mexico, and increasing international capital mobility.<sup>3</sup>

Despite the clear theoretical presumption and strong anecdotal evidence, the magnitude of the effect of unions on establishment or firm survival is uncertain. One limiting factor is the absence of large, representative data sets that track establishments over time and provide information on union status.<sup>4</sup> A second and even more important concern is the fact that unionization is nonrandom. Depending on the correlation between factors associated with higher or lower risks of survival, and higher or lower likelihood

---

<sup>1</sup> See Lewis [1986].

<sup>2</sup> The presence of a "deadweight welfare loss" to unionization is a staple of textbook treatments of unionization. Even Freeman and Medoff [1984] - whose emphasis is on the possible "efficiency enhancing" aspect of trade unionism - stipulate the existence of such a welfare loss. More recently, the cross-country analysis of Nickell and Layard [1999] suggest that a change from 25 to over 70 percent of the workers covered by collective bargaining is associated with a doubling of the unemployment rate. Lalonde, Marschke, and Troske [1996], using a "difference-in-difference" approach with LRD data, find successful organization is associated with significant declines in subsequent employment and output.

<sup>3</sup> The Commission for Labor Cooperation, a tri-national organization created under the North American Agreement on Labor Cooperation ("NAALC") in response to labor issues related to NAFTA, called for a study on the impact of plant closings on union organizing in the three countries.

<sup>4</sup> This has led researchers to use creative data collection methods to examine these questions. For example, Freeman and Kleiner [1990] conducted on-site interviews of 364 establishments that experienced representation elections in the Boston and Kansas City NLRB districts. Bronars and Deere [1993] construct a dataset of NLRB elections to COMPUSTAT data to construct a panel of 85 firms over a 20-year period. Freeman and Kleiner [1999] also use COMPUSTAT to construct a sample of 319 firms. Lalonde, Marschke, and Troske [1996] match NLRB representation elections to a subset of manufacturing establishments that are continuously operating in the LRD to create samples with 500 to about 1100 observations.

of unionization, the observed correlation between union status and survival may overstate or understate the true effects of unions. Two competing phenomena may induce opposite selectivity biases. On the one hand, unions may tend to organize at highly profitable enterprises that are more likely to survive. On the other, a union organizing drive may be more likely to succeed when a firm is poorly managed, or has faced recent difficulties.

In this paper, we present quasi-experimental evidence on the causal effect of unionization on the probability of business failures/re-locations, using a new database that is representative of U.S. establishments at risk of being unionized. Our analysis is based on the fact that most new unionization occurs as a result of a secret ballot election. By law, a majority vote in favor of the union requires management to recognize the union and bargain “in good faith”. This process creates a natural set of comparisons between establishments that faced elections where the union *barely* won (say, by one vote) and those that faced elections where the union *barely* lost (by one vote). As in other regression discontinuity designs, the comparison between near winners and near losers eliminates any confounding selection and omitted variable biases, and allows us to devise credible and transparent estimates of the effect of unions on firm survival.<sup>5</sup>

Our analysis yields a surprising result: we estimate a very small union recognition effect on employer survival, on the order of 1 percent. The estimates are stable across different time horizons (from 1- to 18-years), and across industry, and different establishment sizes. Moreover, we find strong empirical support for the basic requirement of our research design. Close winners and losers look similar in industry composition, size of the bargaining unit, and other pre-determined characteristics.

A potential concern is that unions that barely lose elections nonetheless have a relatively good chance of eventually becoming recognized, while those that barely win face a high risk of decertification. The “treatment” of a *close* union victory would have little long-run consequence to the firm, which would naturally lead to small long-run effects on business survival. We find no empirical support for this hypothesis. There is a striking discontinuity in the relationship between the union vote share and our independent

---

<sup>5</sup> Regression discontinuity designs are described in Thistlethwaite and Campbell [1960] and Campbell [1969], and formally examined as an identification strategy recently in Hahn, Todd, and van der Klaauw [2001]. More recent examples include Angrist and Lavy [1998], and van der Klaauw [1996].

proxy for union presence (the filing of a contract expiration notice) many years after the election.

One explanation for the small effect of unions on firm survival is that the costs imposed by newly organized unions are small.<sup>6</sup> This possibility would seem at odds with the extensive literature on union wage premiums and other research showing significant effects on profitability.<sup>7</sup> It would also be at odds with the fact that NLRB representation elections are nearly always contentious. Why would employers spend resources on resisting an organizing drive if unions did not impose significantly higher future costs? And why would unions spend resources on organizing if there were nothing to gain from a victory? A simple model of optimal union organizing actually suggests that these potential gains should be *larger* when the vote is close than when the union is certain to prevail. A larger wage gain is needed to offset the lower probability of winning the election, in order to justify the fixed costs of an organizing drive. All of this suggests that the costs of new unionization are likely to be important.

We therefore conclude that unions likely do not affect businesses by making them more susceptible to failure or re-location, despite the fears of many employers and employees. While not affecting the survivability of a firm, unions could nonetheless cause slower employment growth. Our data provide limited evidence on this “within-firm” effect: we estimate a negative employment response of 7 percent among surviving establishments, but with standard errors of the same magnitude. Our estimates on employment can rule out the large magnitudes that would be needed to fully explain the decline in union density since the early 1980s. On the other hand, even with the more than 27,000 establishments in our data set, we cannot rule out either small or moderate-sized elasticities.

The paper is organized as follows. Section 2 describes our empirical strategy, and summarizes our main findings. Section 3 discusses the conditions for a valid regression discontinuity design in our context, and outlines a simple economic model to describe the interaction of employers, unions, and workers in the context of a representation election. Section 4 places our analysis in the context of industrial relations in the U.S., and describes some important aspects of our data. We present the main results in Section 5, discuss

---

<sup>6</sup> Freeman and Kleiner [1990] find in their survey of 364 firms, quite modest wage effects associated with new unionization, in contrast to that implied by the typical study based on micro data on individuals.

<sup>7</sup> For example, see Ruback and Zimmerman [1984], and Abowd [1989].

their economic implications in Section 6, and suggest directions for future research in Section 7.

## 2 Regression Discontinuity Analysis of Establishment Survival and Unionization: Basic Facts

We begin by briefly describing our research design and summarizing our main results. Our regression discontinuity analysis suggests that there is virtually no *causal* impact of unionization on survival rates of business establishments.

The primary challenge of identifying the impact of unionization is one of isolating exogenous variation in the union “status” of an establishment, while keeping all other pre-determined characteristics of the establishment “constant”. A simple comparison of survival rates between unionized and non-unionized firms suffers from the potential endogeneity of union status. For example, firms with larger economic rents may be more likely to survive, and *because* of those rents, are more likely to generate worker demand for a union.

In this paper, we exploit a distinctive feature of the union recognition process in the U.S. that we argue generates “as good as randomized” variation in union status. In the U.S. the obligation of employers to bargain “in good faith” with a union is almost always determined through a majority secret-ballot vote amongst the workers; this “representation election” is overseen by the National Labor Relations Board (NLRB). We argue that it is plausible that there is at least some degree of unpredictability regarding the eventual vote tally, so that the firms and unions involved in elections where the union *barely won* (say, by one vote) are likely to be *ex ante* comparable to those firms and unions involved in elections where the union *barely lost* (by one vote). If this is true – that they are comparable “in all other ways” – then the difference in the survival rates subsequent to the election can be attributed to a causal impact of the union certification.

Figure 1a graphically summarizes our main empirical finding. Using data on NLRB representation elections conducted between 1983 and 1999, it plots the proportion of employers that are still in business by the year 2001, by the actual share of the votes in favor of the union. Each dot represents an average

among elections within a 0.05 interval of the vote share. All points to the left of the 0.50 line represent establishments where the union lost the certification election, and those to the right represent union victories. If unionization substantially impacted ability of the employer to remain economically viable – a claim frequently touted by employers during organizing drives [Bronfenbrenner 1994, Kleiner 2001] – we would expect to see a sharp drop in survival rates at the 0.50 threshold. No drop is evident in Figure Ia.

*We interpret this smooth empirical relation through the 0.50 threshold as evidence that there is little or no causal effect of union status on employer survival.* Two alternative explanations to the pattern in Figure Ia are 1) employers in which the union *barely* won and lost are *not ex ante* comparable, perhaps because there is *no unpredictability* in the outcome of the election, and that there is “sorting” of unions and employers on either side of the 0.50 threshold and 2) the *intensity* of the “treatment” is small among the employers and unions involved in close elections – that is, an employer that prevents union certification by 1 vote must make wage concessions as large as they would have had the union won by 1 vote.

As we will discuss in greater detail in later sections, there are several reasons why the two interpretations are difficult to reconcile with the implications of a simple economic model of the election process, as well as with the empirical evidence we present. First, if the exact vote tally could be predicted with complete certainty, it is difficult to explain why a union would expend resources to ultimately lose an organizing drive. At the very least, if vote tallies could be perfectly predicted for a substantial *fraction* of the election cases (and if there were systematic “sorting” around the 0.50 threshold) we would expect to see a sharp drop in the relative frequency of observed elections in which the union barely lost an election. In fact, as Appendix Table II and Appendix Figure I show, union losses are at least as common as union victories, and the distribution of vote shares looks approximately normal, centered around 0.40, with no sharp drop in the relative frequency of bare union losses.

More importantly, we present evidence below that employers and elections, in fact, look quite similar along many pre-determined characteristics on either side of the 0.50 threshold, as predicted by our implicit assumption of “as good as random” assignment. As one example, Figure Ic plots averages of our “union presence before the election” proxy – an indicator of whether a contract expiration notice

was filed prior to the election – against the election vote share for the union.<sup>8</sup> The figure reveals no sharp change in pre-election “union presence” around the 0.50 percent threshold, suggesting that at least along this dimension, employers involved in close wins and losses appear comparable.

Furthermore, the economic model we outline below illustrates that while the expected union wage could be expected to be higher given a higher probability that the union will win, we might also expect employers to offer a higher *non-union* wage in order to persuade some workers to vote against the union. Therefore, the *gap* between the expected union and non-union wage may not necessarily rise with the probability of a union victory. Actually, the *minimum* union-non-union wage gap needed in order to justify the unions’ costs of conducting an organizing drive must be *larger* for elections in which there is a *lower* probability of a union victory. Put simply, it would be unlikely that an election would be conducted if the union had nothing to gain from winning (i.e. if the expected “treatment” is zero).

More importantly, we present evidence that is inconsistent with the hypothesis that union certification among close elections is ineffectual in altering union power. Figure 1b is analogous to Figure 1c, except we plot averages of our proxy for “union presence” *subsequent* to the NLRB representation election. It exhibits a striking discontinuity at the 0.50 threshold, suggesting that the certification has a significant causal influence in altering union power. Thus, it is evident in the data that something *is* “at stake” among these election cases.

The remainder of the paper explains in greater detail the above reasons why we interpret our findings as evidence of little or no *causal* impact of unionization on employer survivorship.

### **3 Econometric Framework**

#### **3.1 Reduced-Form Regression Discontinuity Framework**

In the context of union representation elections, the internal validity of our regression discontinuity analysis primarily depends on two assumptions: 1) that the “treatment” (union recognition) is a known, discontinu-

---

<sup>8</sup> The sample for the figure includes only those employers that have survived as of the year 2001. Below we discuss how to interpret these graphs in light of sample selection bias.

ous function of an observed variable (the votes for the union), and 2) that there is at least *some* unpredictable component of the exact vote tally that has a smooth continuous distribution. U.S. labor law ensures 1); the importance of 2) is discussed below within a reduced-form econometric framework. We outline sufficient stochastic assumptions for identifying union effects in our context.

Suppose an employer outcome, such as the propensity of the establishment to survive, is determined by the equation

$$y_j^* = X_j\gamma + WIN_j\beta + \varepsilon_j \quad (1)$$

where  $y_j^*$  is the survival propensity, and  $WIN_j$  is an indicator variable determining whether the union won the representation election for employer  $j$ .  $X_j$  denotes all other pre-determined characteristics (observable and unobservable) that influence establishment survival and  $\varepsilon_j$  an independent error term.

It is clear that a “regression” of  $y_j^*$  on  $WIN_j$  will in general yield an inconsistent estimate of  $\beta$  if any components of  $X_j$  also determine  $WIN_j$ . As one example,  $X_j$  could represent the magnitude of the firm’s economic rents, and it is plausible that unions are more likely to prevail in organizing drives in firms that enjoy high rents, because workers are less concerned that a union-induced wage gain will harm the economic viability of the employer. In particular, the NLRB election process implies that  $WIN_j$  is determined by

$$\begin{aligned} WIN_j &= \begin{cases} 1 & \text{if } V_j > \frac{1}{2} \\ 0 & \text{if } V_j \leq \frac{1}{2} \end{cases} \\ V_j &= v(X_j) + U_j \end{aligned} \quad (2)$$

where  $V_j$  is the vote share for the union, and  $v(\cdot)$  is a function of  $X_j$ .<sup>9</sup>

The endogeneity of  $X_j$  is clear here. Another way of viewing the problem is that the distribution of  $X_j$ , conditional on a union win, will in general be different from the distribution conditional on a union loss. More formally, (assuming that  $X_j$  has finite and discrete support)  $\Pr [X_j = x | V_j > \frac{1}{2}]$  will generally not equal  $\Pr [X_j = x | V_j \leq \frac{1}{2}]$  for any  $x$ .

---

<sup>9</sup> To be more precise, one could let  $V_j = \Lambda(v(X_j) + U_j)$  where  $\Lambda$  is a transformation from the real line to the  $[0, 1]$  interval.

On the other hand, if  $U_j$  is continuously distributed, the distribution of  $X_j$  conditional on a *bare* union victory will be arbitrarily close to the distribution conditional on a *bare* union loss. That is,

$$\Pr \left[ X_j = x | V_j = \frac{1}{2} + \Delta \right] \approx \Pr \left[ X_j = x | V_j = \frac{1}{2} - \Delta \right] \quad (3)$$

for all  $x$  in the support of  $X_j$ , as  $\Delta$  approaches zero.<sup>10</sup> This is the sense in which examining closer and closer elections can result in employers involved in union victories and losses becoming more similar along all other observable and unobservable characteristics. Thus, a test of the continuity of the distribution of  $U_j$  is to examine if elections appear different along observable pre-determined characteristics of employers involved in elections just below and above the 0.50 threshold.

In practice, rather than specifying functional forms for the distributions of  $X_j\gamma$  and  $\varepsilon_j$  to estimate a normalized  $\beta$ , we will directly estimate  $\Pr [Survival|V_j = \frac{1}{2} + \Delta] - \Pr [Survival|V_j = \frac{1}{2} - \Delta]$  (the discontinuity “gap” at 0.50 illustrated in Figures Ia-Ic). It is clear that a negative (positive) difference implies a negative (positive) impact  $\beta$  of unionization on employer survival propensities.

### 3.2 Conceptual Framework

While our data are not sufficiently rich to allow estimation of a dynamic structural model of union and firm behavior, we present a simple economic model provides a framework for interpreting our estimates and for describing the election context. Drawing upon ideas in Rosen (1969), we present a simple economic framework that describes the first-order objectives and constraints of the various participants in the NLRB election process.<sup>11</sup> Below, we show that the key identifying assumption (continuity of the distribution of  $U_j$  at  $V_j = \frac{1}{2}$ ) is a natural consequence of this simple economic model of the election process.

Consider the structural relation between employer survival propensities and the wage  $W_j$

$$y_j^* = \alpha W_j + \varepsilon_j \quad (4)$$

<sup>10</sup> To see this, note that the density of  $v(X_j) + U_j$  conditional on  $X_j = x$  –  $f(v|X_j = x)$  – is continuous in  $v$ . Therefore, by Bayes’ Law,  $\frac{f(v|X_j=x) \cdot \Pr[X_j=x]}{\sum f(v|X_j=x) \cdot \Pr[X_j=x]} = \Pr[X_j = x | v(x) + U_j = v]$  is continuous in  $v$ . Actually, all that is needed is continuity in  $v$  at  $v = \frac{1}{2}$ .

<sup>11</sup> The model presented below has a similarity to those used to model final offer arbitration [see Farber, 1980; Ashenfelter and Bloom 1984]. Here, the union and management put forth wage “offers” optimally, given expectations of the preferences of the voters/employees, which are unknown to the two bargaining parties. The outcome of the election is akin to the decision process of an “arbitrator”.

$\alpha$  is assumed to be negative, so that a higher wage leads to a higher probability that the employer is forced to shut down. Note however, that even if  $\alpha$  is negative, given the necessarily non-linear relation between a probability and  $W_j$ , it could very well be that the marginal effect of  $W_j$  on the *probability* of survival could be small.

As in Ashenfelter and Johnson (1969), we assume that there are three separate agents, 1) the employer, 2) the union leadership, and 3) the workers – the voters – in the potential bargaining unit. Management and the union each offer different levels of wages to the workers, and the workers vote for or against the union, based on those choices, so that we have

$$\begin{aligned} W_j &= W_j^U WIN_j + W_j^N (1 - WIN_j) \\ WIN_j &= 1 \left( V_j > \frac{1}{2} \right) \end{aligned} \quad (5)$$

where  $W_j^U$  and  $W_j^N$  are the wages offered by the union and management, respectively. This implies the relation

$$y_j^* = \alpha W_j^N + \alpha (W_j^U - W_j^N) WIN_j + \varepsilon_j \quad (6)$$

so that the causal effect of the union is represented by  $\alpha (W_j^U - W_j^N)$ .

**Workers** When considering whether or not to vote in favor of the union, the worker weighs the benefit of gaining a higher wage against the costs of a potentially higher probability that she will not retain employment at the firm. A lower probability of retaining the job may arise either because the union will induce the employer to shut down or move, or induce the employer to scale back employment. Indeed, a worker may find the wage (and its consequences) offered by the management to be more reasonable and individually desirable. We describe the aggregate voting behavior of the workers as

$$V_j = v(W_j^U, W_j^N) + U_j \quad (7)$$

where  $v(\cdot)$  translates the offered wages into a predictable component of the ultimate vote tally, and  $U_j$  is the unpredictable (by all agents in the model) component of the union vote share.<sup>12</sup>

<sup>12</sup> Again, we could be more precise instead specifying  $G^{-1}(V_j) = v(W_j^U, W_j^N) + U_j$ , where  $G$  is a one-to-one transformation from the real line to the  $[0, 1]$  interval.

The shape of  $v(\cdot)$  characterizes the workers' preferences. We assume that  $\frac{\partial v}{\partial W_j^U}$  and  $\frac{\partial v}{\partial W_j^N}$  are both negative, capturing the notion that if the union raises the offer, it will lose the voters who are indifferent between the union's and management's offers. Similarly, the management can gain more "no" votes by promising a higher wage  $W_j^N$ .

**Employer** The management seeks to maximize profits, weighing the benefits of offering a lower wage against the cost that lowering the wage induces a higher probability that the union will win, resulting in a higher wage  $W_j^U$ . The firm's optimal choice of  $W_j^N$  can be expressed as

$$W_j^{N*} = \arg \max_{W_j^N} H(W_j^N, W_j^U, \Pr[WIN_j = 1]) \quad (8)$$

where  $H(\cdot)$  is the employer's objective function. An example of a specific form for the objective function is the expected profits, given the employer's and union's wage offers:

$$H(W_j^N, W_j^U, \Pr[WIN_j = 1]) = \pi(W_j^N) \Pr[WIN_j = 0] + \pi(W_j^U) \Pr[WIN_j = 1] \quad (9)$$

where  $\pi(\cdot)$  is the profit function.

**Union** The union leadership seeks to raise wages above that offered by the employer, but by offering higher wages, it reduces the probability of winning the election. The union's optimal choice of  $W_j^U$  facing the firm's offer  $W_j^N$  can be written as

$$W_j^{U*} = \arg \max_{W_j^U} J(W_j^U, W_j^N, \Pr[WIN_j = 1]) \quad (10)$$

where  $J(\cdot)$  is the union's objective function. An example of a possible form for this objective function is the expected net benefit (expressed in dollars) for the union:

$$J(W_j^U, W_j^N, \Pr[WIN_j = 1]) = -c + \Pr[WIN_j = 1] U(W_j^U - W_j^N) \quad (11)$$

where  $c$  is a fixed cost to conducting an organizing drive, and  $U(\cdot)$  is a benefit function with  $U(0) = 0$ .<sup>13</sup>

In a Nash equilibrium, in anticipation of how the workers will vote (on average), and given the correct expectation of the other party's wage offer, union and management optimally choose their wage

<sup>13</sup> In principle,  $U(\cdot)$  could be decreasing in the wage gap at some point, if the union also gives weight to the employer's survival and/or the level of employment. However, it is easy to imagine that in many election cases  $U(\cdot)$  is increasing in the wage gap in equilibrium.

offers to maximize their objective functions.

**Implications** It is clear from Equation 11 that if the outcomes of elections were known *ex ante* with certainty – if there were no  $U_j$  component – we should expect to observe no elections in which unions lose, if conducting an organizing drive is costly. In fact, as shown in Appendix Table II, over the sample period, the average vote share for the union and win rates are around 0.48 and 0.427, respectively. Furthermore, if in a large fraction of election cases, the outcome of a potential election were perfectly foreseen, we would expect to see a sharp drop in the relative frequency of a “close” union election loss. Appendix Figure I shows that empirically there is no such sharp drop; the distribution of vote shares is approximately normal, centered around 0.35 to 0.40. Our framework provides an intuitive explanation for these patterns: for every election, there is *some* ex ante probability (however small) that the union will prevail. In other words, there is some unpredictability in the final vote count. This would explain why unions would participate in elections that they so happen to eventually lose.

Furthermore, presuming the existence of an unpredictable component  $U_j$  to the ultimate vote share, the framework above implies that the distribution of  $U_j$  would be continuous at  $V_j = v(W_j^{U*}, W_j^{N*}) + U_j = \frac{1}{2}$ . This is because a *discontinuity* in the distribution of  $U_j$  at  $V_j = \frac{1}{2}$  is *inconsistent* with the optimality of  $W_j^{U*}$  and  $W_j^{N*}$  for the union and management. This is because such a discontinuity would imply that  $\Pr[WIN_j = 1]$  would be discontinuous in  $W_j^U$  and  $W_j^N$  at  $W_j^{U*}, W_j^{N*}$ .<sup>14</sup> As evident from Equations 9 and 11, if this were the case,  $W_j^{U*}$  ( $W_j^{N*}$ ) would not be optimal, since the union (management) could lower (raise) wages by an arbitrarily small amount to affect a sharp rise (fall) in the probability of a union victory.<sup>15</sup> Intuitively, no firm or union would settle with making a wage offer that could be altered a tiny amount in order to cause a discontinuous increase in the probability that the election would result in their favor. Thus, the existence of Nash Equilibria gives a theoretical justification of the validity of the regression discontinuity approach described above.

<sup>14</sup> To see this, note that  $\Pr[WIN_j = 1] = F(v(W_j^{U*}, W_j^{N*}) - \frac{1}{2})$ , where  $F(\cdot)$  is the cdf of  $-U_j$ . A discontinuity in the distribution of  $-U_j$  at  $V_j = \frac{1}{2}$ , implies that  $F(\cdot)$  is discontinuous at  $v(W_j^{U*}, W_j^{N*}) - \frac{1}{2}$ . As long as  $v(W_j^{U*}, W_j^{N*})$  is a continuous function of its arguments at  $W_j^{U*}, W_j^{N*}$ , then  $\Pr[WIN_j = 1]$  is discontinuous in  $W_j^U, W_j^N$  at  $W_j^{U*}, W_j^{N*}$ .

<sup>15</sup> This is true as long as the profit function  $\pi(\cdot)$  and benefit function  $U(\cdot)$  are continuous.

Nevertheless, we reiterate that the internal validity of the design can be empirically evaluated. It is a testable proposition whether or not the examination of close elections yields “treated” and “control” employers that look comparable along the observable dimensions of  $X_j$ .<sup>16</sup>

## 4 Background

### 4.1 Institutional Background: the industrial relations climate and the NLRB Election Process

Our sample of establishments is limited to the employers that are at risk for becoming “unionized.” Thus, before proceeding any further, it will be instructive to place our analysis in the context of labor relations and the conduct of representation elections in the U.S.

The administration of fair, secret ballot elections to determine union recognition is one of the chief responsibilities of the NLRB, the most significant administrative agency to be a consequence of the National Labor Relations Act (NLRA) – the Wagner Act – of the 1930s. The law has been changing continuously since its enactment, most notably with the passage of Taft–Hartley Acts in 1947 (which among other things, provided for temporary government seizure of struck facilities in the event of a strike that creates an “emergency”) and the Landrum–Griffin Act of 1959 (which among other things, outlawed a number of successful union tactics including “secondary boycotts”). In principle, the NLRA provides a neutral setting in which the right for workers to bargain collectively is enforced.

It is important to note that where U.S. law gives workers the right to unionize, an NLRB representation election is not required. In general, nothing prevents an employer from recognizing a union without the formalities of an election. Voluntary recognition of a union, however, is thought to be quite rare, and employers generally will attempt to resist an organizing drive. With data on firms who faced NLRB elections in the early 1990s, Brofenbrenner (1994) documents that most employers used multiple tactics to delay or deny a collective bargaining agreement. Among the most common are

---

<sup>16</sup> As in any empirical investigation, it is impossible to rule out violation of the identifying assumption, but we *can* test the restriction of the research design that *observable* characteristics are roughly balanced between the union-win and union-loss groups of employers.

1. "Captive meetings". While employers are prohibited from directly firing workers because of lawful union activity, at captive meetings employers are allowed to inform workers of the possible (dire) consequences of unionization, including making the business more susceptible to closure.
2. Firing union activists. While "prohibited," the penalty imposed on employers, if found guilty, is generally quite minor – reinstatement with back pay. Indeed, the costs have been perceived as so minor that Freeman (1985) observes that the notices that firms are required to post when they engage in illegal firing are referred to as "hunting licenses."
3. Hire a "management consultant" who advises employers on a variety of tactics to discourage unionization.<sup>17</sup>
4. Alleging unfair labor practices, disputing the choice of bargaining unit, etc.<sup>18</sup>

Against this backdrop of employer opposition to unionization, it is perhaps not surprising that there is no single path to an NLRB election and eventual recognition of the union by the employer. Nonetheless, it is useful to describe a prototypical scenario that results in a establishment agreeing to bargain with its workers through a labor union:

1. A group of workers decide to try to form a union. These workers contact a labor union and ask for assistance in beginning an organizing drive.
2. In collaboration with the union, the employees begin a "card drive." The purpose of the card drive is to be able to petition the NLRB to hold an election. Unions generally seek to get cards from at least 50 percent of the workers in the 6 month period of time usually allowed (although in principle, only 30% is required to be granted an election by the NLRB.)
3. After the cards have been submitted, the NLRB makes a ruling on whether the people the union seeks to represent have a "community of interest" – basically form a coherent group for the purposes of bargaining. The NLRB makes a determination of which categories of employees fall within the union's "bargaining unit." Often the parties will differ on the appropriate bargaining unit – employers generally prefer larger and more heterogeneous groupings than do unions.
4. Next, the NLRB holds an election at the work site (with exceptions to account for such things as the vagaries of employment seasonality). A simple majority (50 percent plus 1 vote) for one union is all that is required to win.<sup>19</sup>
5. Within 7 days after the final tally of the ballots, parties can file objections to how the election was conducted. In principle, with sufficient evidence that the election was not carried out properly, the NLRB can rule to invalidate the outcome of an election, and conduct another one thereafter. Specific ballots cannot be challenged after the voting is completed.
6. If after this, a union still has a simple majority, then the employer is, in principle, obligated to negotiate "in good faith."

Two aspects of the industrial relations climate deem our sample of establishments particularly ap-

---

<sup>17</sup> For a colorful, albeit idiosyncratic discussion see Levitt (1993).

<sup>18</sup> In the case of graduate students at universities, for example, employers have often attempted to argue – sometimes successfully – that graduate student employees are not "employees" but "students receiving financial aid." Another example is employers arguing that its employees are not workers but "independent contractors" who are not covered by the provisions of NLRA.

<sup>19</sup> If two unions split the vote 50-50, they both lose, and neither become certified.

appropriate for an analysis of the impact of unionization on employer outcomes.

First, that employers are thought to generally oppose organization drives [Kleiner 2001] suggests that both parties have “something at stake” in the outcome of the election. For example, we expect that both the union and management are expecting that a union win will generally lead to higher wages, more benefits, or better working conditions, at the cost of the employer. If very little were at stake, and if the elections themselves were *pro forma* events, then we would not expect to see a significant employer response to a union election victory. Such a finding would say more about the small size of the “treatment” (“unionization”) than the potential magnitude of distortions caused by an aggressive union. This seems unlikely, however, since we analyze establishments which faced NLRB elections. Such a focus would seem likely to select establishments where union-management relations are contentious since in the overwhelming majority of cases the management of such establishments always has the option of *voluntarily* recognizing the union *without* a (costly) NLRB election. Thus, it would seem more reasonable to assume that the outcome of the election is far from inconsequential to both parties.

Second, combined with a contentious atmosphere, the secret-ballot nature of the vote undoubtedly generates a certain amount of uncertainty in the outcome of the election, particularly when the vote is expected to be close. As shown in Section 3 a certain degree of uncertainty is a crucial element to our theoretical and econometric framework. The assumption of some randomness to the vote would likely not be justified if union certification could be secured through a public petition. For example, if all that was required were 50 percent or more signatures, one could imagine that the sample of establishments/unions where the unions submitted a petition with 51 percent of the signatures would be very different from a (peculiar) group of establishments/unions where the workers submitted signatures that totalled 49 percent. By contrast, it is very easy to imagine in a secret-ballot context that those unions that won 26 out of 50 votes in a secret-ballot election possessed virtually the same *ex ante* chance of winning as unions that obtained 25 out of 50 votes (and lost).

## **4.2 Data Set Construction: the NLRB, FMCS, and InfoUSA, Inc.**

Deferring the details to the Appendix, we summarize here the most important features of the dataset used in the analysis. First, electronic records on all representation election cases handled by the NLRB in the fiscal years from 1984 to 1999 were obtained. These records have information such as the dates of the filing of the petition, the election, and the closing of the case, as well as the eventual vote tallies, as well as other characteristics such as the size of the voting unit, and the primary industry of the establishment in question.

Importantly, these files contain the establishment name and exact address. The names and addresses alone were submitted to a commercial marketing database company called InfoUSA, Inc. InfoUSA maintains an annually updated list of all active business establishments (with a telephone listing) in the United States. The basis for their database is the consolidation of virtually all telephone books in the country. InfoUSA makes a brief call to each establishment at least once a year, to verify their existence, and to update their information on various items such as 1) the total number of employees at the establishment, 2) the estimated sales volume of the establishment, 3) the primary product of the business, and various other characteristics. If InfoUSA found a record in their current database (as of May, 2001) with the same name and address, they appended their information to the record. They were not given any information beyond the name of the business and the street address.

This merged data was then additionally merged to a database of all contract expiration notices between 1984 and February, 2001 – more than 500,000 case records – obtained from the Federal Mediation and Conciliation Service (FMCS) through a Freedom of Information Act (FOIA) request. According to the U.S. Code of Federal Regulations (29 CFR 1425.2)

In order that the Service may provide assistance to the parties, the party initiating negotiations shall file a notice with the FMCS Notice Processing Unit ... at least 30 days prior to the expiration or modification date of an existing agreement, or 30 days prior to the reopener date of an existing agreement...

Thus, in principle, parties to collective bargaining agreements are required to file so-called “30-day notices” with the FMCS. This was used to obtain our proxy for the “presence” of a union at the establishment – both before and after the election – under the presumption that contracts eventually expire,

typically after two or three years.

There are a few important limitations to our data. First, our data do not constitute a true panel dataset. We only observe “survival” or “death” as of one point in time - in the year 2001. We know little about what happens between the time of the election and 2001, except the observation of contract expirations at that particular location. While we do observe a few “baseline” characteristics from the NLRB election file, InfoUSA does not retain historical records, so we do not have detailed employment and sales data for period between the election and the year we observe survival status (2001).

Second, since we are measuring employer “survival” as a match (by name and address) in the InfoUSA database, there will undoubtedly be some measurement error. Consequently, we will inevitably treat some firms as having “died”, when instead we have simply been unable to match them. However, while this may mean that estimates of the *level* of survival rates may be downward biased, it is highly unlikely that establishments with close union winners are systematically less or more likely to match to the InfoUSA database than counterpart close union losers, except if there is a true impact of union certification on survival probabilities.

Likewise, our measure of “union presence” will also likely be biased in *levels*, although this is unlikely to have important consequences for our comparison of close winners to close losers. For example, we understate the extent of unionization to the extent that our matching algorithm fails to locate a match in the FMCS data when such a match exists or to the extent that noncompliance with the law (regarding notifying the FMCS when a contract expires) is widespread. Alternatively, we overstate the extent of unionization to the extent that our matching algorithm produces “false positive” matches. Although the levels may be mis-measured, it seems reasonable to assume that these sources of measurement error are unlikely to be systematically different between close winners and close losers. On balance, we believe the benefits of being able to compare the bare winners and losers on the basis of some other measure of “union presence” other than the legal certification that *necessarily* results from winning the election outweighs the inability to obtain an accurate measure of the overall *level* of union presence.

### 4.3 Descriptive Statistics

Since the primary outcome of interest in this analysis is the survival probability of the employer, we first provide a broad picture of what our data implies about exit rates of establishments over time. Three important patterns emerge from our sample of establishments that experience NLRB representation elections: 1) as might be expected, establishments' survival probabilities decline as one examines longer and longer intervals, 2) employers where the organizing campaign succeeded have, on average, a lower probability of surviving than those where the union was defeated, and 3) establishments' death/exit rates appear to be the dominant component to the overall decline in the total employment they provide over time. Table I illustrates these basic patterns.

Table I and the subsequent analysis in this paper is essentially based on the universe of establishments that experienced NLRB representation elections between 1983 and 1999.<sup>20</sup> The first column of Table I shows that in this sample, as would be expected, the probability of survival declines significantly as one examines longer and longer intervals. For example, the table shows that among the employers that experienced NLRB representation elections in 1984, roughly 28 percent of them were still in existence as of the year 2001. By contrast, about 58 percent of establishments that experienced an election in 1999 had survived as of 2001. The survival probability grows monotonically as we examine more recent elections; this would be expected if the establishments that experienced elections were, on average, comparable over time. The implied exit rates are comparable to other estimates from existing research that utilizes establishment-level longitudinal data.<sup>21</sup>

The next three columns of Table I show that establishments where unions were recognized are, on average, less likely to survive both in the short- and long-run. For example, among elections that occurred in 1999, 54 percent of those employers where the union won were still "alive" by the year 2001, compared

---

<sup>20</sup> The NLRB election data are representation election cases that are disposed within the fiscal years from 1984 to 1999; thus, most of the elections were held between the years of 1983 and 1999, with a few elections occurring before 1983.

<sup>21</sup> For example, Dunne, Roberts, and Samuelson [1989] report a 5-year exit rate of about 40 percent, as calculated from Census of Manufacturing data (1967-1977). And an analysis of food-manufacturing plants by McGuckin, Nguyen, and Reznick [1998] imply a 10-year exit rate of about 60 percent in the LRD from 1977 to 1987. As seen in Table I, the corresponding implied exit rates in our data are 48 and 59 percent respectively.

to 61 percent for the non-unionized employers. There is also a difference in survival rates 13 years after the election (elections in 1988). Averaging across all years yields a statistically significant difference of 3 percentage points. Presumably, this difference reflects a combination of both the *causal* impact of union recognition on survival, and the likelihood that there are systematic differences between employers that won or lost the election.

Taken together, the fifth and ninth columns in Table I suggest that in this sample, employer death or exit is a significant component in the decline of total employment provided by employers over time.<sup>22</sup> Among employers that faced elections in 1984, the average employment level - where “dead” employers are counted as having zero employment - is about 61, while the corresponding numbers for elections conducted in the late 90s are over 100.<sup>23</sup> By contrast, the average log employment (ninth column) *conditional on survival by 2001* appears to be relatively stable over time. This implies that much of the decline over time in the average employment provided by establishments operates through the “deaths” of employers.<sup>24</sup>

Finally, the differences between the 10th and 11th columns show that, conditional on survival, employment levels are much smaller among employers where the union was victorious. Again, this difference (an average of 0.22 in logs over the entire sample period) presumably reflects a combination of both the causal impact of unionization on employment levels, and a selection bias resulting from systematic differences between employers that do and do not successfully resist an organizing campaign.

A more detailed examination of the differences between employers where the union won or lost the election reveals three additional patterns: 1) establishments where unions won are significantly smaller, in terms of sales volume as measured in 2001, than those where unions lost, 2) across establishments, the election outcome is significantly associated with our own proxy of union presence constructed from the FMCS data, and 3) the election outcome is also associated with several other pre-determined characteristics of the establishment (e.g. industry, size of the voting unit).

---

<sup>22</sup> The employment variable is employment level, as of the year 2001. The variable comes from the InfoUSA, Inc. database.

<sup>23</sup> All “dead” establishments were assigned zero employment. There are 17622-16355=1267 missing values for employment among the surviving employers as of 2001.

<sup>24</sup> Longitudinal evidence indicates that establishment “deaths” alone can account for a significant share of job destruction in the manufacturing sector. On average, 25 percent of job destruction in the manufacturing sector is due solely to establishment “deaths” [Davis and Haltiwanger 1992]. The rate appears to be higher in other sectors of the economy [Pivetz, Searson, and Spletzer 2001]

Table II provides the details of these findings. The fifth row reports that employers where the union wins produce roughly 35 percent less in sales compared to cases where the union lost. Row (6) in Table II also shows that the outcome of the representation election is highly correlated with our proxy for union presence. This computation is made among the restricted sample of surviving (as of 2001) establishments. Among “union-loss” establishments we observe a contract expiring a approximately 10 percent of the time. When the union wins the election, on the other hand, there is a 36 percent chance that we will observe a union contract ending after the election.

The rest of Table II provides good reason for the analyst to resist interpreting the union-won/loss differences in rows (1) - (5) as causal effects of union certification. For example, the establishments where the union won are about 15 to 20 percent smaller than the “union-loss” establishments, as measured by the number of eligible voters or the ultimate number of votes cast in the NLRB election. In light of these differences, it is thus not surprising that we observe differences in employment *after* the election, in the same direction, and of roughly the same magnitude.

Similarly, row (7) of Table II reveals that we are more likely to detect the presence of a union *before the election* among establishments where the union won recognition compared to employers where the union lost: the respective proportions of pre-election “union presence” are 0.179 and 0.095.

In addition, employers differ by election outcomes on a number of other dimensions; these differences give more reason to maintain some doubt in any causal interpretation of the comparisons in rows (1) - (5) of Table II. For example, as row (12) of Table II indicates, establishments where the union won the election are much less likely (33 versus 42 percent) to be classified in the manufacturing sector. Moreover, as rows (13) and (14) of Table II indicate, establishments where the union won are more likely to be in the service sector (35 percent in “union-win” establishments versus 22 percent for “union-loss” establishments) and the voting unit less likely to be classified by the NLRB as “truck drivers”. On the other hand, measures of state economic conditions are not strongly related to the outcome of the election. The union won/loss differences in the levels and changes in the unemployment rate and the log(employment) level are statistically but not economically significant.

In sum, Table II provides evidence that *caution the analyst against* making inferences about the impacts of union certification on employer outcomes from simple differences in outcomes by election outcome.<sup>25</sup> The evidence is suggestive that union recognition by election may be *negatively selected* – that unions are more likely to prevail in a representation election in smaller, and potentially *less* robust establishments. This would be consistent with the notion that larger establishments with greater resources may be more able to resist organizing drives. However, this interpretation of Table II is at best speculative without an independent estimate of the causal effect of union certification.

## 5 Estimates of the Impact of Unionization

### 5.1 Evidence on Validity of the Regression Discontinuity Design

As mentioned above, if the regression discontinuity design is valid, employers involved in close union wins should be similar, on average, to those involved in close union losses, in terms of their pre-determined characteristics – whether or not they are observed by the econometrician. As in any empirical analysis, assessing whether “unobservables” are balanced is impossible. However, we can at least assess whether or not the regression discontinuity design is succeeding in balancing *observable* determinants of establishment survival. Our empirical analysis reveals that the examination of close elections *does* result in “treatment” and “control” establishments that appear to be otherwise similar on observable dimensions.

Table III reports differences in characteristics of the employer, by union victory/loss, in the overall sample, and by sub-samples that isolate closer and closer elections. For example, in elections where the union lost, the average number of eligible voters is about 113, compared to about 92 where the union eventually won. That difference remains large when we examine elections where the union won between 25 and 75 percent of that vote. But it drops in half when we focus on the comparison among elections where the union won between 35 and 65 percent of the vote. The same holds true in percentage terms. The difference in terms of the log of the vote cast falls from -0.19 to about -0.10 when we move from the first

---

<sup>25</sup> The researcher might be tempted to conduct the analysis conditional on the pre-determined characteristics such as industry, and size of voting unit, under the presumption that the election outcome is random conditional on those covariates. Besides being somewhat ad hoc, by “using up” the covariates, this approach has the drawback of eliminating any possibility of gauging the internal validity of the comparison.

to third set of columns.

The differences in these average characteristics become even smaller when we focus our attention on all elections where the share of the vote for the union is between 0.45 and 0.50. For twenty-person votes, this means the outcome was decided by one vote. Table III shows that, for example, the differences in the pre-election size of the employer (as measured by the number of eligible voters and votes cast) fall to 2 or 3 (on bases of more than 100). Employers involved in union losses are more likely to be classified as manufacturing establishments in the overall sample (first set of columns), but that difference falls to a statistically insignificant  $-.025$  when examining elections decided by the narrowest margin (the fourth set of columns). The monotonically decreasing differences, as one compares closer and closer elections, is also true for the proportion of employers categorized as service sector establishments, and for the proportion of voting units classified by the NLRB as “trucking”.

Table III also shows that as one examines closer elections, the estimated standard errors rise, which would be expected, since the number of observations used for the analysis necessarily declines. At some point, restricting the sample to even closer elections will result in *no* observations for analysis. This illustrates the well-known trade-off between bias and variance in non-parametric estimation of an unknown conditional expectation function. Indeed, the averages in the last set of columns can be interpreted as kernel regression estimates using a uniform kernel of bandwidth 0.05. Insofar as the slope of the true conditional expectation function (of the variables with respect to the vote share) is nonzero, *any* kernel regression estimate will necessarily be biased in finite samples.

A simple alternative way of using data points away from the 0.50 threshold to estimate population means at the threshold, is to specify a flexible-form parametrization of the underlying function. Figures IIa, b, and c illustrate the results from regressing the corresponding dependent variables on a fourth-order polynomial in the vote share, including a dummy variable  $V_j > 0.50$ , and the dummy variable interacted with a linear term in  $V_j$ . The predicted values of those regressions are super-imposed upon local averages by 0.05 intervals.

Figure IIa, IIb, and IIc reveal 1) that the predicted values from the polynomial come reasonably

close to the local averages, 2) there is a generally smooth empirical relation between, for example  $\log(\text{total votes cast})$ , the probability that the employer is a manufacturing or service sector establishment, and the observed vote share, and 3), there is no striking discontinuity in any of these relations at the 0.50 threshold.

Overall, Table III and Figures IIA-c lend credence to the interpretation of any discontinuity in *post-election* outcomes as the causal effect of union recognition.

## 5.2 Effects on Survival Rates

The survival rate differences between union victories and losses, by the margin of victory/loss, are reported in the first row of Table III. It shows that the modest difference in survival probabilities in the overall sample falls to a point estimate of -0.023 in the second set of columns. It falls further to -0.014 in the third set of columns, and when the sample is restricted to elections where the union vote share is between 0.45 and 0.55, the difference becomes a statistically insignificant, -0.007 with a estimated standard error of 0.015. Given the magnitude of the standard error, the estimate is consistent with both a negative effect of up to -.037, as well as a positive impact of up to 0.023 at conventional levels of significance.

Similar to Figures IIA-c, Figure Ia reveals a relatively smooth relation between the fraction of surviving employers and the vote share for the union, with no apparent discontinuity at the 0.50 threshold. Again, the polynomial specification (the same specification used in Figures IIA-c) yields predicted values that closely track the local averages by vote share. The polynomial specification also yields a causal estimate of around -0.007 with a standard error of 0.014.

**Robustness to Alternative Specifications** Our causal estimates of the impact of unionization of employer survival are robust to a variety of different specifications. The robustness of the estimates to varying specifications is in fact an implication of a valid regression discontinuity design. Intuitively, *if examining close elections generates “as good as randomized” variation in union status, then not only should union status be independent of pre-determined “baseline” characteristics (Table III), but the inclusion of those characteristics as covariates in a regression analysis should not significantly alter the point estimates of the effect of certification.*

Table IV provides evidence that the estimates are indeed robust to various specifications. The estimate that results from the basic polynomial specification (depicted in Figure 1a) is reported in Column (1) (top panel) of Table IV. The first row of Table IV demonstrates that the estimate of the impact of union recognition on survival probabilities is robust to the inclusion of any combination of pre-determined characteristics, with the estimates ranging from -0.003 to -0.007. In Column (2), a set of election-year dummy variables is included. In Column(3) state dummies are additionally included, and in Column (4) industry dummies and unit dummies (indicating how the NLRB classified the primary type of worker in the bargaining unit) are included. In Columns (5) and (6) our proxy for the presence of a union *before* the election is included (since it is pre-determined as of the election date), as is the log of the number of eligible voters in the bargaining unit. It is important to note that the significant coefficient on the pre-election union presence dummy should not be interpreted as a *causal effect*, but simply indicative that this proxy appears to absorb some of the “residual” variance in the dependent variable.

We consider one final robustness check. If the regression discontinuity design is valid, we can replace the survival indicator with a “residualized” version of the indicator (the residuals from a regression of the survival indicator on all the covariates) and find a similar estimate. Indeed, when we use the resulting residuals as the dependent variable in a polynomial specification in the vote share (as in Column (1)), Column (7) shows that the estimate is -0.007 – nearly identical to the other estimates in the first row.

**Estimates of Heterogeneous Effects** We next examine the extent to which the “treatment” effect of unionization, as discussed above, varies along three potentially important ways. By analogy to Ordinary Least Squares with a binary treatment variable and other covariates, we examine whether there any interesting “interactions” between the treatment and other exogenous covariates. As the results from Table V make clear, however, the data cannot reject the hypothesis that the estimated “treatment” effect of unionization, as presented above, are constant along the three dimensions we explore.

First, since we observe survival at one point in time (the year 2001), the “overall” estimate we obtained is an average of 2- through 18-year survival rates. Since a *positive* effect on survival in the short-

run could be canceling out a long-run *negative* effect (or vice versa), it is instructive to stratify the analysis by groups of years. The 2nd row of the first column presents the estimate from Column (6) of Table IV. The 3rd row reports that among elections that were held before 1988, the corresponding effect of union recognition on the probability of survival is about -0.022 with a standard error of 0.020. The following three rows report the interaction effects for the periods 1988-1991, 1992-1995, and after 1995, respectively. The interactions effects are small and statistically insignificant.<sup>26</sup>

Second, the effect could potentially vary significantly by industry. The next three rows in the first column of Table V show that the effects are slightly larger for service sector establishments, but again the interactions are not jointly statistically significant. Third, the effects could vary by the size of the voting unit (a rough proxy for initial size of the establishment). The final three rows of Table V show that the estimates are positive (0.005) for voting unit sizes between 20 and 40 workers, and are slightly negative (0.005-0.032=-0.027) for units between 40 and 100. But all interactions are not statistically meaningful at conventional levels of significance.

The second column reports that the estimates of the overall effect and the various interactions do not change significantly, when we use an alternative measure of establishment survival. As mentioned in Section 4, we consider that an establishment has survived as of 2001 if the company name and address matches an entry in the InfoUSA database. However, in principle, if bare losers are much more likely than bare winners to undergo an ownership change – and hence change their name – then our primary measure of establishment survival may mask a true effect on establishment closure. A comparatively robust way to address this issue is to consider that an establishment has survived if *any* establishment is present at that exact address as of the year 2001 – irrespective of whether the company name changes. The second column of Table V reports the estimates from using this measure (Survival (2)), and shows that the estimates mirror that of the first column.<sup>27</sup>

On balance, while we cannot rule out small heterogeneous effects by these three observable di-

---

<sup>26</sup> Specifically, the specification was the same as Column (6), except year dummies were replaced with time-period dummies, their interactions with the  $WIN_j$  indicator and their interactions with  $V_t$ .

<sup>27</sup> Apparent from the first row, the obvious exception, as expected, is that the proportion of establishments that have *any* business (regardless of name) as of the year 2001, is significantly higher, at about 0.643.

mensions due to the magnitudes of the sampling errors, we interpret the estimates as indicating that our main estimate is not being wholly driven by a particular sub-sample, as defined along these observable dimensions.

### **5.3 Further Evidence on the Consequences of Union Recognition**

It is tempting to conjecture that a union that barely wins a representation election possesses the same degree of “bargaining power” as a union that barely lost. After all, even if the NLRA mandates that the employer bargain “in good faith” with the union as the exclusive representative of the workers in the unit, there is no guarantee that a first contract will be secured. Also, since nothing prevents a losing union from attempting another organizing drive (as long as one year has passed since the first election), one might conjecture that the outcome of a *close election* may in the long-run have no effect on the extent of union presence or power at the employer.

As we discuss in greater detail in Section 6, this conjecture is difficult to reconcile with the simple economic framework we have outlined above. Among other reasons, it is difficult to explain why a union would undergo a costly organizing campaign if the outcome of the election was inconsequential for the union.

More importantly, this conjecture is strongly inconsistent with our analysis of the data we have collected on post-election union presence. Specifically, while we do not directly observe the securing of a first contract, we observe the expiration of a first contract (and expirations of subsequent contracts). Thus, our post-election “union presence” variable is whether or not we observe at the employer’s location an FMCS contract expiration notice subsequent to the date of the election. The second to last row of Table III shows that there is a large difference in the probability of observing an expiration notice between union victories and losses. That difference becomes smaller, when examining closer elections, but remains a highly statistically significant 0.159 when examining elections with a union vote share between 0.45 and 0.55. In fact, our post-election “union presence” variable is the only variable in Table III where the differences remain large when examining the closest of elections.

The behavior of our post-election “union presence” proxy also stands out in our graphical analysis of the data. Figure Ib shows a smooth empirical relation between the vote share and the probability of observing a contract expiration notice – everywhere except at the 0.50 threshold, where there is a sharp jump in that probability from about 0.12 to 0.28. By contrast, the last row of Table III and Figure Ic, suggest no corresponding striking change for our pre-election “union presence” variable (an indicator of whether we observe a contract expiration notice at the address of the employer *before* the election). These figures suggest that the outcome of the representation election is far from inconsequential – that there is a permanent *causal* effect of a union victory on this particular proxy of union power. Also, the effect on our post-election union presence proxy is robust to alternative specifications, as shown in Table IV. This would be expected if the assumptions of the regression discontinuity design were valid. The estimates of the effect of union recognition on the probability that we observe a contract expiration notice at the establishment are precisely estimated and range from 0.132 to 0.152. Note from the comparison of Columns (3) and (5) or (4) and (6) that the inclusion of the *pre-election* union proxy does not meaningfully affect the union recognition effect, despite its own independent predictive power (t-statistic over 20).<sup>28</sup>

Some care needs to be taken in interpreting these findings in Table III and Figures Ib and Ic, because the averages are computed using the sample of establishments that have survived as of the year 2001. In principle, this induces a censored sample selection problem. In the same way that wages are not observed for the non-employed, our post-election union presence variable – and any other measure of union presence – is defined only for those surviving establishments<sup>29</sup>, there is therefore the potential that the discontinuity in Figure Ib is an artifact of sample selection bias [Heckman 1976].

However, in this particular context, the sample selection bias problem may not be important after all. Lee [2002] shows that if a treatment is “as good as randomized”, and if treatment affects sample selection in a monotonic way, then equal probabilities of selection in the treatment and control group imply that there is *no* sample selection bias. These two conditions are applicable here: 1) the maintained hypothesis

---

<sup>28</sup> Again, the coefficient on the pre-election union presence variable is not to be interpreted as a causal effect; rather it is more properly thought of as a partial correlation; its inclusion “absorbs” residual variation.

<sup>29</sup> Actually, we do have some information for those that are not in business as of the year 2001, but we do not know when, between the date of the election and 2001, the establishment shut down.

is that the regression discontinuity design generates “as good as randomized” variation in the union status treatment, and the data fail to reject the restrictions implied by that hypothesis, and 2) Equations 1 and 5 both satisfy the monotonicity condition in Lee [2002]. Given that we estimate little or no effect of union recognition on survival (here, sample selection), this suggests that there is little or no sample selection bias in our causal estimate of the effect of union recognition on post-election union presence.

Furthermore, Lee [2002] shows that if the selection probabilities are the same, the independence and monotonicity conditions imply that the *selected sample* of treated and control populations should possess similar distributions of baseline (or pre-determined) characteristics. Such a pattern is apparent in these data, as shown in Appendix Table III, which is analogous to Table III, except that it restricts the sample to establishments that have survived as of the year 2001.<sup>30</sup>

Finally, since the evidence suggests that sample selection bias may not be important, we proceed by examining union impacts on four other measures of employer outcomes: the levels of employment and estimated sales volume, as well as the logs of both variables. Table V reports regression discontinuity estimates for these outcomes. Overall, the results are mixed; the employment and sales responses are small relative to the overall variability in outcomes across establishments. However, the estimated standard errors are too large to rule out economically meaningful negative or positive effects on employment.

The third column reports regression discontinuity estimates for the level of employment, where we have assigned “0” to those establishments that have closed by the year 2001. The point estimate is essentially 0 with a standard error of 8. While the null hypothesis of a zero effect cannot be rejected at conventional levels of significance, neither can a negative response of about 13 employees – about 15 percent of the mean (83.4).

The fifth column reports the estimates for estimated sales volume, where again we have assigned “0” to establishments no longer in existence. The estimates imply a positive impact on sales volume on the order of \$250,000 dollars, but it is not statistically different from zero. Furthermore, a substantial

---

<sup>30</sup> Technically, it also conditions on having non-missing employment information. There are some surviving establishments that did not report employment levels.

negative effect (of about \$2.6 million, about 20 percent of the mean) could not be ruled out at conventional levels of significance. As with employment, the cross-sectional variability in sales is quite significant, and suggests that even large, economically meaningful effects on sales are small, relative to the cross-sectional heterogeneity in output across establishments, even while including industry dummies.

The fourth and sixth columns report the discontinuity estimates for the log of employment and sales volume, respectively; hence, “dead” or “0 employment” establishments are necessarily dropped from the sample. The null hypothesis that the effects are zero cannot be rejected at conventional levels of significance, but the estimated standard errors are themselves of economically significant magnitudes. Nonetheless, we are able to statistically rule out large effects: for example, a 25 percent negative effect on employment and a 20 percent positive effect on sales volume.

## **6 Economic Implications**

We believe that our estimates constitute an important step in empirically assessing the magnitude of economic distortions caused by unions. Considering it as a first step to providing a broader picture of the causal effects of unions (while adequately considering the potential biases induced by self-selection and omitted variables), we are hesitant to draw strong inferences regarding the magnitude of potential welfare losses. However, it is instructive to consider our findings’ implications in light of five issues of economic interpretation.

### **6.1 Is the “treatment” small?**

An alternative explanation of the lack of effects on survival rates is that the wage demands by unions, on average, are themselves small – too small to have a measurable impact on employer survival rates. It is tempting to conjecture that even though a union election win causes greater union presence (as measured by the existence of a contract), a union that barely wins an election cannot negotiate wages any larger than what would have been offered by the management had the union barely lost.

However, the notion that the gap between the union and management offered wage is small among close elections, and large among votes where the union happens to overwhelming win, is not a robust

prediction of our framework of optimizing employers and unions. Within our framework, it is possible to generate specific examples where such a prediction would hold, given specific assumptions about the shapes of the objective functions for the employer and union, the shape of  $v(\cdot)$ , and the distribution of  $-U_j$ . On the other hand, it is just as easy to generate the opposite prediction – that the gap in the wage offers  $W_j^U - W_j^N$  declines with an exogenous increase in the probability of a union victory.

To see this, suppose that  $\delta_j$  is a parameter that characterizes how much the workers care about wages so that we have  $v(W_j^U, W_j^N, \delta_j)$  where  $\frac{\partial v}{\partial \delta_j} > 0$ . Now suppose that the union leadership has a limited range of credible wage offers; to make the example stark, suppose that it is (correctly) expected by all agents that if the union wins, the wage  $W_j^U$  will be fixed at  $\overline{W_j^U}$ . In this case, the equilibrium is determined solely by the management's counter-offer wage. One can show that there are some reasonable assumptions under which  $\frac{\partial W_j^{N*}}{\partial \delta_j} > 0$  and  $\frac{\partial \Pr[WIN_j=1]}{\partial \delta_j}$ , which would imply that the wage gap  $W_j^{U*} - W_j^{N*}$  diminishes with an exogenous increase in the probability that the union will win.<sup>31</sup> Intuitively, faced with an exogenous increase in wage demands on the part of the workers (increase in  $\delta_j$ ), fearing the prospect of paying the high union wage, management offers a larger non-union wage in order to reduce the chance of losing the election to the union. In equilibrium, the probability of a union victory does increase, but it is moderated by the management's optimal response. Therefore, in this case, we would expect to see a larger wage difference among close elections (where the probability of a union victory is more moderate) compared to elections in which the union overwhelmingly won (in which the probability was higher).

There is another reason to expect that  $W_j^{U*} - W_j^{N*}$  may even be larger with a lower probability of the union winning: a union requires a larger minimum gain  $W_j^{U*} - W_j^{N*}$  in order to justify a costly

<sup>31</sup> Let  $v(W_j^U, W_j^N, \delta_j) = -W_j^U - W_j^N + \delta_j$ . Then the first-order conditions for the employer amounts to the equation

$$\frac{f(v)}{1-F(v)} - \frac{\frac{\partial \pi(W_j^N)}{\partial W_j^N}}{\pi(W_j^U) - \pi(W_j^N)} = MB(v) - MC(W_j^N) = 0, \text{ where } f(\cdot) \text{ and } F(\cdot) \text{ are the density and cdf for } -U_j. \text{ The first term can be interpreted as the marginal benefit from raising the offered wage, which reduces the probability that the union will win, and the second term can be thought of as the marginal costs (due to lower profits). An interior solution is guaranteed if } \frac{f(v)}{1-F(v)}$$

and  $\frac{\partial \pi(W_j^N)}{\partial W_j^N}$  are decreasing and increasing, respectively, in  $W_j^N$ . This will be true if  $-U_j$  is normally distributed and

$$\frac{\partial^2 \pi(W_j^N)}{(\partial W_j^N)^2} < \frac{\left(\frac{\partial \pi(W_j^N)}{\partial W_j^N}\right)^2}{\pi(W_j^N) - \pi(W_j^U)}. \text{ By the implicit function theorem } \frac{\partial W_j^N}{\partial \delta_j} = \frac{\frac{\partial MB}{\partial v}}{\frac{\partial MB}{\partial v} + \frac{\partial MC}{\partial W_j^N}} \text{ which will be positive but less than 1.}$$

Therefore, an increase in  $\delta_j$  leads to an increase in the offered wage  $W_j^N$  and in the equilibrium probability that the union will win.

organizing drive. This is immediately apparent from Equation 11, which shows that the expected benefit – itself a product of the probability and the benefit function  $U(\cdot)$  – needs to exceed  $c$  to justify having an election in the first place. Put simply, the lower probability that the union will prevail, the larger wage gain is needed to justify the costs of organizing.

For these reasons, it seems more plausible that the outcome of the election is far from inconsequential to both the employer and the union, perhaps *especially* among close elections (where the *ex ante* probabilities of a union win are more moderate). However, whether or not the outcome of a representation election has any measurable impact on compensation and working conditions is ultimately an empirical question. Future research could show that the outcomes of elections *are* inconsequential in terms of altering compensation and working conditions, which would suggest that the conventional wisdom regarding how unions affect wages would need to be re-evaluated.

## 6.2 Is the effect “economically” small?

Our analysis obtains estimates around -0.01, and the null hypothesis of “no effect” cannot be rejected at conventional levels of statistical significance. However, it is also instructive to consider whether or not the estimates are consistent with an alternative null hypothesis – *that the decline in the union sector in the U.S. in recent years is entirely attributable to the union impact on employer survival.*

Such “back-of-the-envelope” calculations require a great deal of abstraction. However, suppose we consider an economy initially made up of  $N_0$  establishments with identical, constant hazard rate of closure of  $d$  per year, and a constant inflow of  $bN_0$  establishments per year. Normalizing  $N_0 = 1$ , it is straightforward to show that the number of establishments at time  $t$  is represented by

$$n(t) = \lambda + (1 - \lambda)e^{-dt} \tag{12}$$

where  $\lambda = \frac{b}{d}$ , so that if  $b = d$ , the number of establishments remains constant over time.<sup>32</sup>

Over the period from 1983 to 1998, union density among private sector workers fell by approximately 40 percent. The implied baseline hazard rate from our data of establishments implies a constant

---

<sup>32</sup> This follows from a the differential equation  $n'(t) = dn(t) + b$ .

hazard of  $d = 0.10$ .<sup>33</sup> This implies  $\lambda = (0.6 - e^{-0.10(15)}) \frac{1}{1 - e^{-0.10(15)}} = 0.485$ . Thus, if the decline of the union sector was entirely due to the union effect on the ability of establishments to survive (i.e. if the establishments were not unionized,  $\lambda$  would be 1), then the causal effect of union recognition would have to be a doubling of the hazard rate  $d$ . With a base hazard rate of 0.10, this implies that we should expect causal estimates of the 1-15 year survival rates to be on the order of -0.20.<sup>34</sup> To the contrary, however, our causal estimates are around -0.01 and we are able to reject magnitudes larger than -0.04.

Thus, our estimates cast considerable doubt on the proposition that the primary mechanism of recent union decline is through union effects on establishment closure. Of course, these rough calculations should be viewed with caution, as they are based on a number of compositional, aggregation, and behavioral assumptions – the empirical relevance of which can only be assessed with large-scale micro-data on establishments in the U.S.

### 6.3 The Impacts of Union Threat

Our estimates of the impacts of unionization are *conditional* on having an NLRB election, and our model makes it clear that we are identifying the impacts of  $W_j^U - W_j^N$  on survival rates. As our model makes explicit, however, there is another potential impact of unions, which is the distortion due to “union threat” [Rosen 1969]. In other words, employers may face a higher closure rate because their own offered wage  $W_j^N$  is forced to be significantly higher than the competitive market wage  $W_j^M$ . There are two components to the “overall” impacts of unionization: one component is the effect that arises when wages are pushed from  $W_j^M$  to  $W_j^N$ , and the other component – which we have estimated in this paper – results when wages are increased from  $W_j^N$  to  $W_j^U$ . Since we find little evidence of effects on survival of the second kind, we conclude that evidence of large union effects on closure would more likely be found in analyses of union threat effects.

However, our theoretical framework cautions against making the conjecture that threat effects

---

<sup>33</sup> Actually, our data seem to fit a Weibull baseline better; however, as usual, it is difficult to know whether the empirical duration dependence is “real” or an artifact of unobserved heterogeneity. The 10 percent hazard rate is also roughly consistent with exit rate estimates from Dunne, Roberts, and Samuelson [1989] and McGuckin, Nguyen, and Reznak [1998].

<sup>34</sup> This can be seen by comparing the survivor function  $e^{-.1t}$  and  $e^{-.2t}$  over a 15-period interval.

should necessarily be larger than the union effects that we have estimated in this analysis. To see this, suppose that the firm will survive only if  $y_j^*$  in Equation 4 exceeds some scalar  $a$ . Then the probability of closure is  $\Pr[Closure] = F(a - \alpha W_j)$ , where  $F(\cdot)$  is the cdf of  $\varepsilon_j$ .  $\frac{d\Pr[Closure]}{dW_j} = -f(a - \alpha W_j) \alpha$  and  $\frac{d^2 \Pr[Closure]}{dW_j^2} = f'(a - \alpha W_j) \alpha^2$ ; the sign of the second derivative depends on the shape of the distribution of the error term, and can easily be positive, which would mean that a marginal impact on closure rises with the wage. Thus, a wage increase that results from a union victory, which is *incremental* on top of an already existing above-market wage due to union threat – could easily lead to a larger effect on closure, compared to the effect induced by an equally sized wage gain due to union threat.

Ultimately, the magnitude of the union threat effect requires empirical investigation.

#### 6.4 Is “no effect” in survival rates consistent with monopoly unionism?

Put simply, yes. There are many margins on which unions could cause allocative distortions. First, even if employers’ probabilities of closure are entirely unaffected, they could cut their level of employment. This could be occurring *within* every establishment in the union sector. Second, even if employment is not affected, if the employer’s production techniques are constrained by the terms of a contract or union threat, productivity and hence output could fall.

It is instructive to consider a benchmark for what one might expect to observe in the analogous case where adjustments occur primarily through the “intensive” margin of employment. Interpret  $\lambda = \frac{(1-\Delta)b}{d} = 1 - \Delta$  in terms of the proportion of total employees at an establishment, so that  $\Delta$  is now the “within-employer” response. If we assume that the inflow of new establishments equals the hazard rate, under the null hypothesis that the sole reason for the union decline was due to an intensive margin effect, we would expect the measured causal impact of union certification on employment to be about a 50 percent decline in the level of employment at the establishment.<sup>35</sup> Even given the imprecision of our intensive margin estimates, however, we can easily reject an effect of such magnitude. On the other hand, we cannot rule out a smaller negative intensive margin employment effect of unionization of, say, 15 percent. Given the

<sup>35</sup> Alternatively, if the effects were on the margin of affecting the appearance of establishments in the union sector, it is easy to show that if it entirely explained recent union decline in the U.S., the effect would be a quite large effect of 0.50 in probability.

often cited estimate of the union wage effect of 15 percent [Lewis, 1963] this would imply a labor demand elasticity of -1.

The magnitude of intensive margin effects is a topic for future research. But examining the extensive margin is a pre-requisite to examining the intensive margin. Even in the ideal scenario (from a purely scientific standpoint) where union recognition were randomly assigned, it would not be straightforward to estimate intensive margin effects on employment if the “treated” employers were more likely to “die” and hence “drop out of the sample”. Our findings that there are negligible survival effects and that among the survivors both the bare winners and losers look similar along observable characteristics; this suggests that the sample selection issue may well be a second order issue.

## 6.5 Is “no effect” unsurprising?

Finally, it may be tempting to argue on purely theoretical grounds that the question of whether or not union recognition impacts employer survival is a non-issue in the first place, if we assume that workers and unions are rational. After all, unionized workers also lose if an employer is forced to close its operations.<sup>36</sup> This is certainly true in the unrealistic case that unionization led to closure with complete certainty. But if the union effect works through a negative impact on the *probability* of survival, a large effect (say a doubling of a hazard rate of 0.10) could easily be consistent with rationally behaving workers and unions.<sup>37</sup> Thus, arguments purely based on the presumed rationality of workers are not informative in bounding the union effects on employer survival.

Nonetheless, the notion that the organizing workers and union leadership may be behaving rationally – and have a direct stake in maximizing the viability of the employer – has important implications. Indeed, not only could the union be interested in reducing the risk of closure, but they could be interested

---

<sup>36</sup> This is pointed out in Kuhn [1986] and Freeman and Kleiner [1999].

<sup>37</sup> If the worker’s outside option is equivalent to the firm’s non-union wage offer, then it is trivial that any worker will prefer to vote for the union, even if there are substantial effects of the wage on closure rates, as long as there is some probability that the establishment will still exist so that they can obtain the union wage.

If the worker’s outside option is below the firm’s offered wage, then for the indifferent worker (under risk neutrality) we would have  $\frac{W_N - W_M}{1 - \delta(1-d)} = \frac{W_U - W_M}{1 - \delta(1-2d)}$ , where  $W_U$ ,  $W_N$ , and  $W_M$  denote the union, firm’s non-union wage offer, and outside options, respectively, and  $\delta$  is the discount factor, and  $d$  is the constant hazard rate for the establishment. If  $W_U - W_N$  were 0.10 in logs and  $\delta$  was 0.95, and  $d$  were 0.10, then  $W_N - W_M$  would have to be about 0.06 (in logs) – a plausible number – in order for the worker to be indifferent.

in maximizing the profitability of the business – insofar as it would lead, for example, to a larger wage bill for the workers. Indeed, this possibility has been the focus of the “efficient contracts” literature (McDonald and Solow (1981), Brown and Ashenfelter (1986), MacCurdy and Pencavel (1986), Card (1986), Abowd (1989), Abowd and Lemieux (1993)). The efficient contracts model departs conceptually from the monopoly union/“right-to-manage” model by observing that union members could value both the level of employment and the wage. And if the firm enjoys some economic rents – through monopoly power in the product market, for example – the monopoly union outcome is “inefficient” in the sense that at least one of the two parties can be made better off without making the other party worse off. In some situations – when there is a “strongly efficient” contract – there is a possibility that unions do not lower employment, but instead act to redistribute rents from firms to workers.

## 7 Conclusion

This study meets two important challenges of credibly estimating the magnitude of the causal effects of unions. First, we have constructed a large data set that represents a virtual universe of establishments facing potential unionization, linked to comprehensive database on survivorship on businesses. Using over 27,000 observations, we have also exploited a feature of the NLRB election process to produce quasi-experimental estimates that are likely to be free of selection and omitted variable biases. Our results suggest that establishment closure is not the main mechanism of the employment reallocation response to unionization. Rather, it seems more likely that employers respond by reducing employment.

In the existing literature, there is a fair amount of variation in estimates of “within-firm” employment responses to unionization.<sup>38</sup> It would seem fruitful to utilize this study’s research design to learn more about this margin of employment adjustment. There is also some uncertainty regarding the labor costs imposed by unions. The existing literature suggests that union wage premiums are large, and that unions have

---

<sup>38</sup> Estimates range from no effect to significant effects. For example, Lalonde, Marschke and Troske [1996] suggest there is significant negative employment response. Card [1990] examines indexed and non-index union contracts and finds a significant negative employment response to higher real wages. On the other hand, Bronars and Deere [1993] examine a panel of publically-traded firms and finds no systematic relation between union growth and firm growth. Freeman and Kleiner [1990] find only modest employment effects (and modest wage effects) associated with new unionization.

significant effects on profitability [Ruback and Zimmerman 1984; Abowd 1989]. At the same time there is some evidence indicating only modest wage gains associated with new unionization [Freeman and Kleiner 1990; Lalonde, Marschke, and Troske 1996]. A useful avenue for future research is to use NLRB election data, as we have here, to explore the impact of unionization on: the level and distribution of wages, capital investment, and total factor productivity.

## **Appendix A.**

There are three sources of information that were merged for the analysis. We first describe how our data from the National Labor Relations Board (NLRB) was matched to firm data from InfoUSA. Next we describe how the result of this matching process, henceforth the “NLRB/InfoUSA data” was matched to data from the Federal Mediation and Conciliation Service (FMCS).

### **A.1 The NLRB to InfoUSA match**

First, electronic records on all representation election cases handled by the NLRB in the fiscal years 1984 to 1999 were obtained. These records contain information such as the dates of the filing of the petition, the election, and the closing of the case, as well as the eventual vote tallies, as well as other characteristics such as the size of the voting unit, and the primary industry of the establishment in question. Most importantly the file contains information on the name and street address at which the representation election was held.

These 139,881 records were then matched by name and address to a commercial marketing database company called InfoUSA, Inc. Before being sent to InfoUSA, however, the address fields were first “standardized” using a program called “Mailers +4 Postal Automation Software.” For example, “1 Broad Street” was changed to “1 BROAD ST”. This was done to facilitate matching the NLRB data to the data from InfoUSA.

As discussed in the text, InfoUSA maintains an annually updated list of all business establishments (with a telephone listing) in the United States. The basis for their database is the consolidation of virtually all telephone books in the country. InfoUSA makes a brief call to each establishment at least once a year, to verify their existence, and to update their information on various items such as 1) the total number of employees at the establishment, 2) the estimated sales volume of the establishment, 3) the primary product of the business, and various other characteristics.

We submitted the name and address information from our “address standardized” NLRB data to InfoUSA who matched as many of the submitted records to their current database (as of May, 2001) and then appended their information to the record. Apart from the name and address information, no other

information was given to InfoUSA.

Before merging this data to our data from the Federal Mediation and Conciliation Service (described below) the data were cleaned for duplicates. There were three types of duplicates: 1) genuine duplicates – more than one NLRB case with a specific employer, 2) duplicates which were an artifact of the fact that our NLRB data came in two files: one contained data from 1977 to 1991 and the other contained the records for 1984-1999. Most of the duplicate pairs therefore occurred for the years 1984–1991, although there were some duplicate pairs in other years because of the fiscal/calendar year distinction; 3) a very small number of duplicates where two records containing exactly the same information.

## **A.2 The NLRB/InfoUSA to FMCS match**

The most difficult task was to take this NLRB and InfoUSA matched data and merge it to a database of 563,565 expiration notices between 1984 and February, 2001 obtained from the Federal Mediation and Conciliation Service (FMCS) through a Freedom of Information Act (FOIA) request. Like the NLRB data the addresses were standardized using “Mailers +4 Postal Automation Software.”

Many of the FMCS addresses could not be standardized. Reasons included:

- The ZIP Code was invalid or does not match the city name.
- Canadian ZIP Code
- Unknown Street. No street could be found within the city that matches the street name given in the address field.
- Non-delivery Address. The location given in the address field physically exists in the city but the post office has indicated that there are no houses there. This may occur when one side of the street has railroad tracks or a river along side it.
- Address Out of Range. The address was not within the delivery range.
- Multiple Addresses. The national database found two or more possible ranges for the address.
- Multiple Components. More than one component of your address was in error. Address components are the pre-direction, post-direction, and suffix.
- No Data Available for City. Address data for this city is not in the database.

We chose to keep all FMCS records that had successfully standardized addresses, and ran an automatic matching algorithm to link NLRB/InfoUSA records to the resulting database. The matching involved matching by address and employer name. Our algorithm first involved removing spaces, punctuation, and

frequently used words in company names (i.e., INCORPORATED, BROTHER, COMPANY, etc.). From this we created two separate strings of a length which we could vary: the first sub-string started from the beginning and the second sub-string begins from the end and went in reverse order. The longer the string, the more “stringent” the test for two records to be declared a “match.” Matches could occur when the first sub-strings matched or when the second sub-strings matched.

Depending on whether the address was standardized, we varied the length of the comparison string – when addresses were standardized then we enforced a less stringent standard than those which were unable to be standardized. Our final algorithm (using only the standardized data) used strings of length 5.

### **A.3 Sample Selection and Variable Construction**

First, in order to minimize measurement error in our survival indicator, we keep only those NLRB records that have a non-missing street address, and those addresses that were successfully standardized using our address standardization software. In addition, we keep all elections in which 20 or more votes were cast. Appendix Table I reports the means and standard deviations of the variables for the entire merged database and the restricted sample used for the estimation. The table shows that the means are reasonably similar between the two samples, except in the variables that reflect the scale of the employer (number of eligible voters, votes cast, employment, sales volume).

Finally, great care was taken to construct the vote share for the union. There is a problem with simply computing the ratio of the number of votes for the union to the total number of votes. This is because there is substantial variability in the number of votes cast, and a union victory is secured by obtaining strictly more than 50 percent of the vote (plus 1 vote). Consider all elections where an even number of votes are cast. Elections with any number of votes cast could result in exactly 50 percent of the vote. However, it is *impossible* that an election with less than 100 votes cast could have a vote share between 50 and 51 percent of the vote. This *mechanically* induces a discontinuity in the size distribution at the 50 percent threshold of the vote share, and this is entirely an artifact of the fact that the vote share is not *literally* continuous, and instead has finite and discrete support, with the support changing with the number of votes cast.

Thus, we made a minor adjustment to the vote share variable in order eliminate this problem. For every case where there was an *even* number of votes cast, an amount equal to  $0.5/(\# \text{ votes cast})$  was subtracted from the vote share. For example, if 25 out of 50 votes were for the union, the vote share became  $0.50 - 0.01 = .49$ ; 26 out of 50 votes meant a vote share of  $0.52 - .01 = 0.51$ . Cases where an odd number of votes cast were unadjusted. This minor adjustment restores symmetry in the support for vote share, and the new “vote share” variable still possesses the property that strictly more than 0.50 implies a union victory. Finally, the vote share was “binned” so that all vote shares between 0.50 and 0.55 were assigned the vote share of 0.525, shares between 0.45 and 0.50 were assigned the share of 0.475, and so forth. In this way, vote shares were standardized to the support for the elections with the smallest number of votes cast (20).

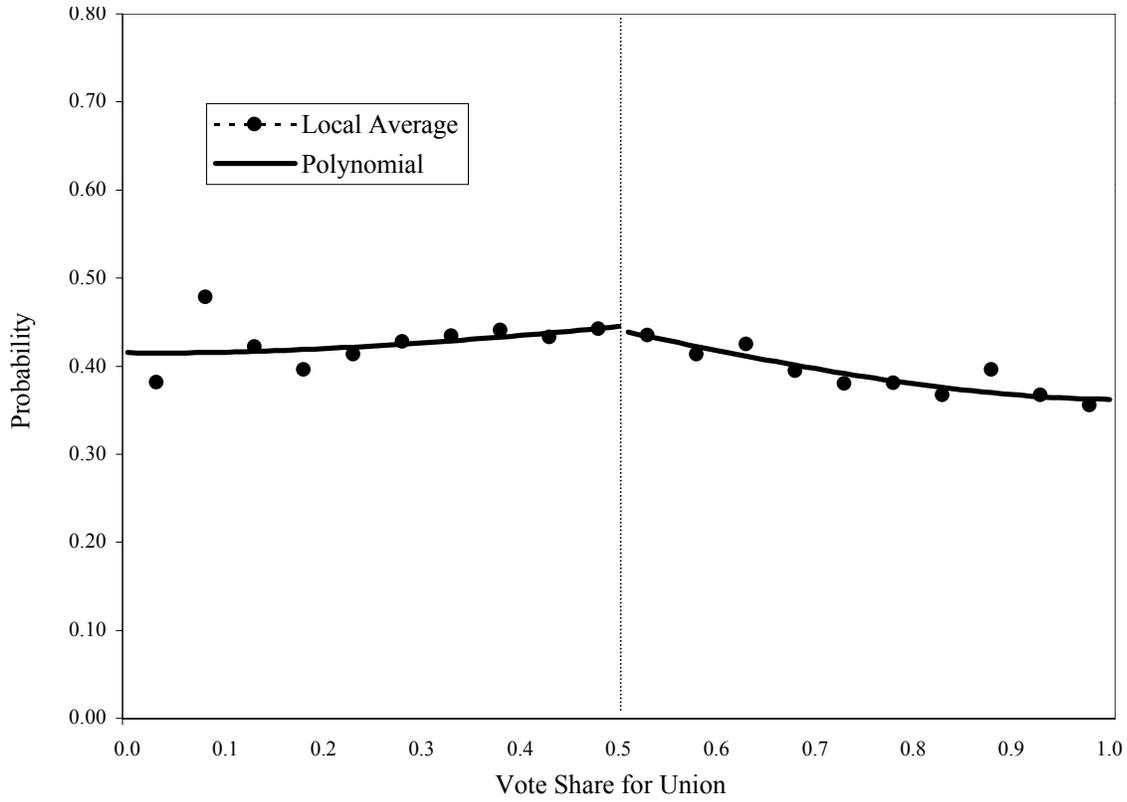
A completely different approach is to abandon the use of the vote share completely, by focusing on the *absolute* vote count, and comparing elections in which the union either won or lost by literally 1 vote. This eliminates this “integer problem”, but at the same time tends to push larger establishments away from the threshold that determines victory (generating a pronounced U-shape in the average size of the establishment, with respect to the absolute vote margin of victory/loss). This was the approach used in DiNardo and Lee [2001]; it should be noted that the results reported there are qualitatively and quantitatively similar to the results in this paper, suggesting that our findings are not sensitive to the method used to address the integer problem.

## References

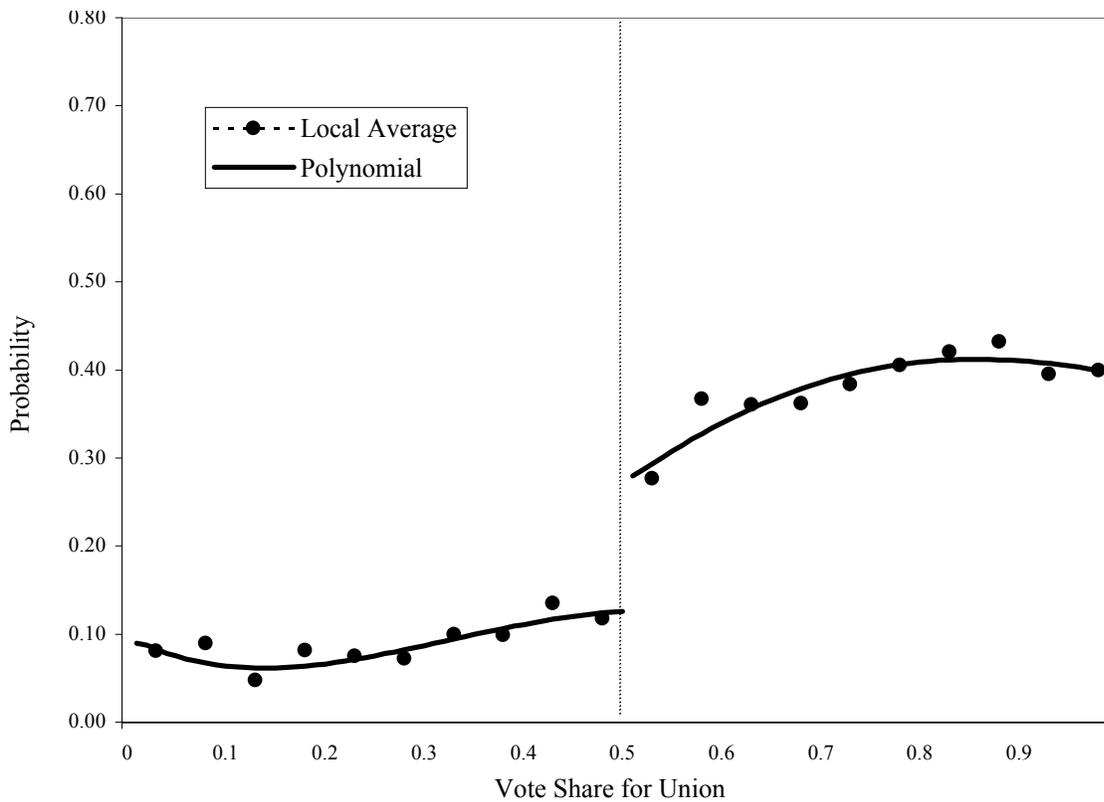
- [1] Abowd, John M. "The effect of wage bargains on the stock market value of the firm" *American Economic Review* 79:774–800, September 1989.
- [2] Abowd, John M. and Thomas Lemieux. "The Effects of Product Market Competition on Collective Bargaining Agreements: The Case of Foreign Competition in Canada." *Quarterly Journal of Economics* 108, (1993): 983-1014.
- [3] Angrist, Joshua D., and Victor Lavy. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114 (1998): 533-75.
- [4] Ashenfelter, Orley and George E. Johnson. "Bargaining Theory, Trade Unions, and Industrial Strike Activity." *American Economic Review*, 59 (1969) 35-49.
- [5] Boal, William M., and John Pencavel, "The Effects of Labor Unions on Employment, Wages, and Days of Operation: Coal Mining in West Virginia," *Quarterly Journal of Economics* 109 (1994) 267-298.
- [6] Bronars, Stephen G., and Donald R. Deere, "Union Organizing Activity, Firm Growth, and the Business Cycle," *American Economic Review*, 83 (1993) 203-220.
- [7] Bronfenbrenner, Kate. "Employer behavior in certification elections and first contracts: Implications for labor law reform" In Sheldon Friedman, Richard Hurd, Rudy Oswald, and Ronald Seeber, editors, *Restoring the Promise of American Labor Law*, pages 75–89. ILR Press, Ithaca, New York, 1994.
- [8] Bronfenbrenner, Kate, *Uneasy Terrain: The Impact of Capital Mobility on Workers, Wages and Union Organizing*, research paper provided by the U.S. Trade Deficit Review Commission, 2000.
- [9] Brown, James N. and Orley Ashenfelter. "Testing the efficiency of employment contracts" *Journal of Political Economy* 94:S40–S87, 1986.
- [10] Campbell, D. T. "Reforms as Experiments." *American Psychologist* 24 (1969): 409-29.
- [11] Card, David. "Efficient contracts with costly adjustment: Short run employment determination for airline mechanics" *American Economic Review* 76:1045–1071, December 1986.
- [12] Card, David, *American Economic Review*, 80 (Sep., 1990), pp. 669-688.
- [13] Commission for Labor Cooperation, *Plant Closings and Labor Rights: A Report to the Council of Ministers on The Effects of Sudden Plant Closings on Freedom of Association and the Right to Organize in Canada, Mexico, and the United States*, 1997.
- [14] Davis, Steven J. and John Haltiwanger, "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *Quarterly Journal of Economics* 107 (1992) 819-863.
- [15] DiNardo, John and David S. Lee. "The Impact of Unionization on Establishment Closure: A Regression Discontinuity Analysis of Representation Elections," *Center for Labor Economics Working Paper #38*, September, 2001.
- [16] Dunne, Timothy, Mark J. Roberts, and Larry Samuelson. "The Growth and Failure of U.S. Manufacturing Plants." *Quarterly Journal of Economics* 104 (1989): 671-698.
- [17] Freeman, Richard B. and James L. Medoff. *What Do Unions Do?* Basic Books, New York, 1984.
- [18] Freeman, Richard, and Morris Kleiner, "The Impact of New Unionization on Wages and Working Conditions," *Journal of Labor Economics*, 8 (1990) S8-S25.
- [19] Freeman, Richard B. and Morris M. Kleiner "Do Unions Make Enterprises Insolvent?," *Industrial and Labor Relations Review*, Vol. 52, No. 4, July 1999, pp. 507- 524.
- [20] Freeman, Richard B. "Why are unions faring poorly in NLRB representation elections?" In Thomas Kochan, editor, *Challenges And Choices Facing American Labor*. MIT Press, Cambridge, 1985.
- [21] Hahn, Jinyong, Petra Todd, and Wilbert van der Klaauw. "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design." *Econometrica* 69 (2001): 201-209.
- [22] Heckman, James. "The Common Structure of Statistical Models of Truncation, Sample Selection, and

- Limited Dependent Variables and A Simple Estimator for Such Models.” *Annals of Economic and Social Measurement* 5 (1976): 475-492.
- [23] Heckman, James J. “Sample selection bias as a specification error.” *Econometrica* 47 (1979): 153-62.
- [24] Kleiner, Morris. “Intensity of Management Resistance: Understanding the Decline of Unionization in the Private Sector” *Journal of Labor Research*, Volume XXII, Number 3, Summer 2001.
- [25] Kuhn, Peter, “Wages, Effort, and Incentive Compatibility in Life-Cycle Employment Contracts.” *Journal of Labor Economics*, 4 (1986) 28-49.
- [26] Lalonde, Robert J., G. Marschke, and K. Troske, “Using Longitudinal Data on Establishments to Analyze the Effects of Union Organizing Campaigns in the United States,” *Annales D’Economie et de Statistique* 41/42 (1996) 155-185.
- [27] Lee, David S. “The Electoral Advantage to Incumbency and Voters’ Valuation of Politicians’ Experience: A Regression Discontinuity Analysis of Close Elections” *Center for Labor Economics Working Paper #31*. UC Berkeley, April, 2001.
- [28] Lee, David S. “Trimming for Bounds on Treatment Effects with Missing Outcomes” *Center for Labor Economics Working Paper #51*, UC Berkeley, March, 2002.
- [29] Lewis, H. Gregg. *Unionism and relative wages in the United States*. University of Chicago Press, Chicago, 1963.
- [30] Lewis, H. Gregg. *Union Relative Wage Effects: A Survey*. University of Chicago Press, Chicago, 1986.
- [31] Macdonald, Ian and Robert M. Solow. “Wage bargaining and employment”. *American Economic Review* 71:886–908, 1981.
- [32] MaCurdy, Thomas and John Pencavel. “Testing the efficiency of employment contracts” *Journal of Political Economy* 94:S3–S39, 1986.
- [33] McGuckin, Robert H., Sang V. Nguyen, and Arnold P. Reznik. “On Measuring the Impact of Ownership Change on Labor: Evidence from U.S. Food-Manufacturing Plant-Level Data.” in *Labor Statistics Measurement Issues*, J. Haltiwanger, M. Manser, and R. Topel, eds. Chicago, Illinois: University of Chicago Press, 1998.
- [34] Nickell, Steven, and Richard Layard, “Labor Market Institutions and Economic Performance,” *Handbook of Labor Economics*, Vol. 3C, Orley Ashenfelter and David Card, ed. Elsevier Science, New York, 1999.
- [35] Pivetz, Timothy R., Searson, Michael, and James R. Spletzer, “Measuring job and establishment flows with BLS Longitudinal Microdata,” *Monthly Labor Review*, April 2001.
- [36] Rosen, Sherwin. “Trade Union Power, Threat Effects and the Extent of Organization.” *Review of Economic Studies* 36 (1969): 185-196.
- [37] Ruback, Richard and Martin Zimmerman, “Unionization and Profitability: Evidence from the Capital Market,” *Journal of Political Economy*, 92 (1984) 1134-57.
- [38] Thistlethwaite, D., and D. Campbell. “Regression -Discontinuity Analysis: An alternative to the ex post facto experiment.” *Journal of Educational Psychology* 51 (1960): 309-17.
- [39] Van der Klaauw, Wilbert. “Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach.” *Unpublished manuscript* (1996).

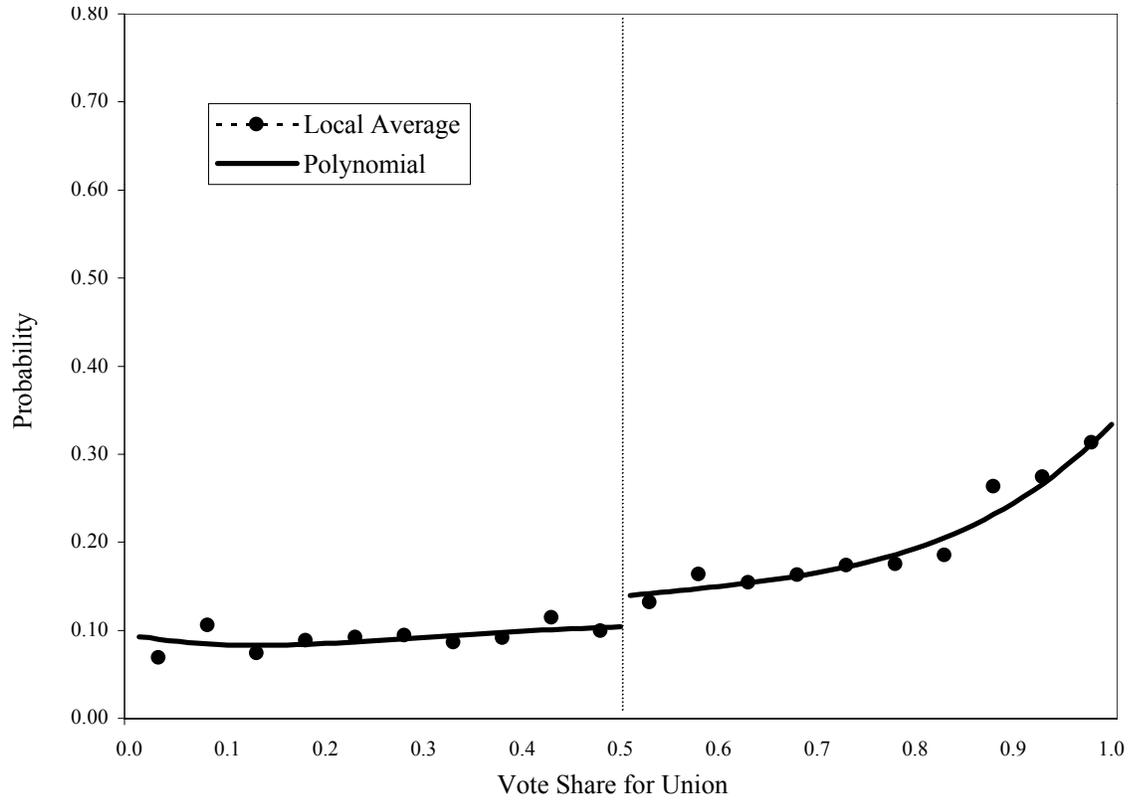
**Figure Ia: Probability of Establishment Survival by 2001, by Vote Share**



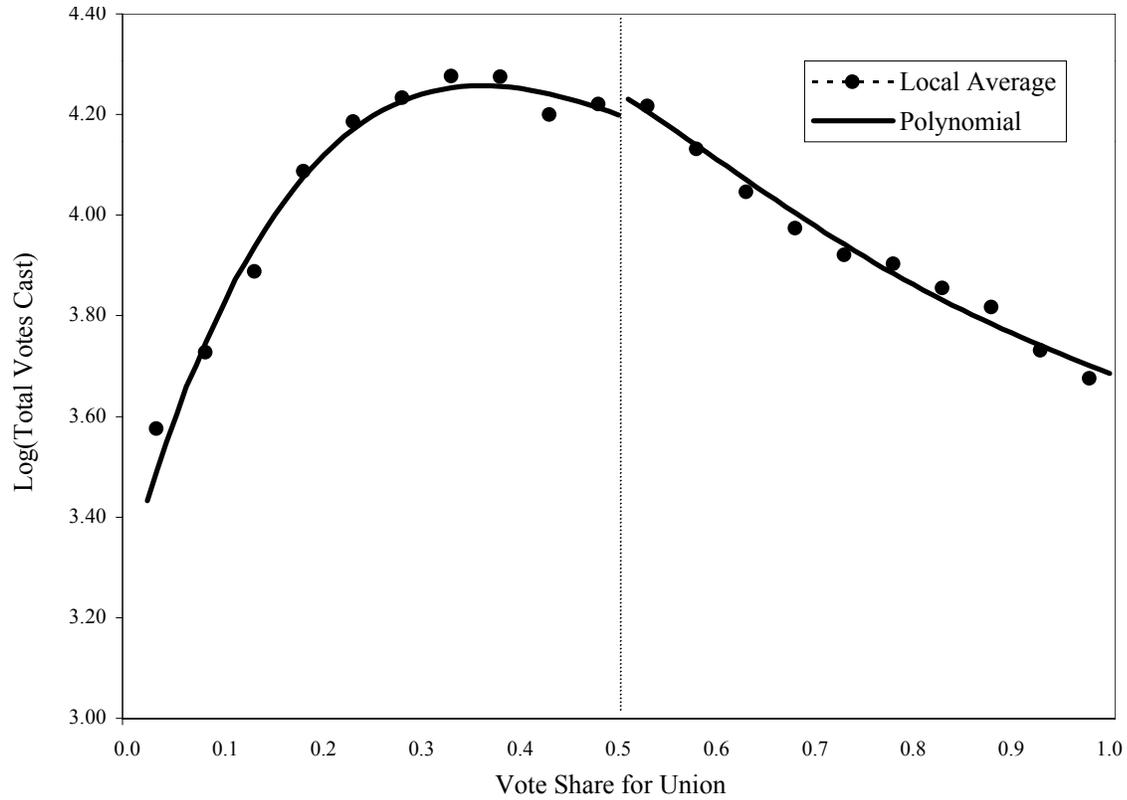
**Figure Ib: Probability of Post-Election Union Presence, by Vote Share**



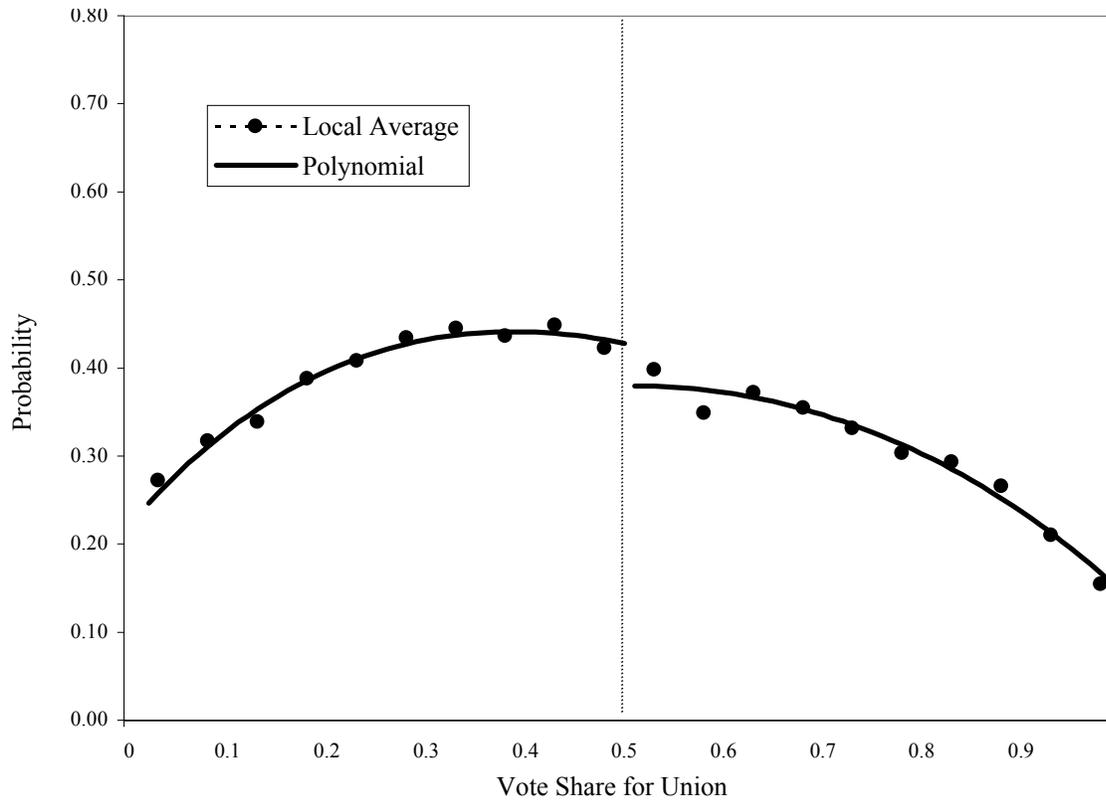
**Figure 1c: Probability of Pre-Election Union Presence, by Vote Share**



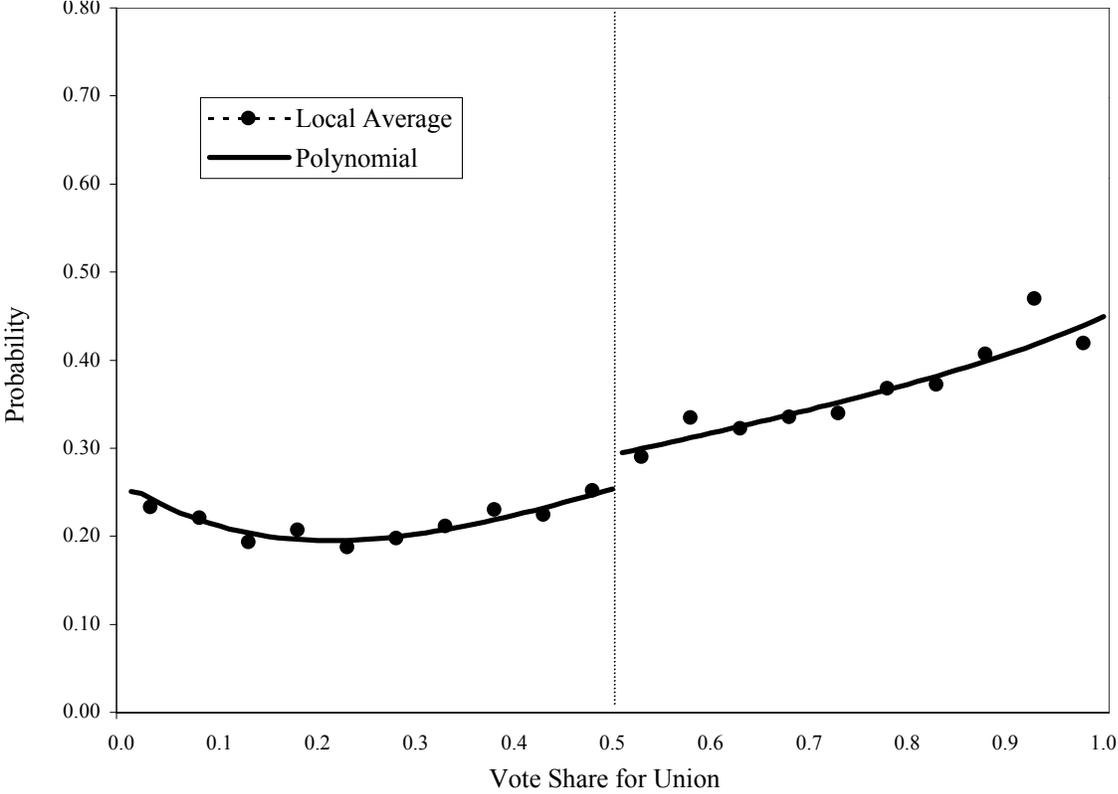
**Figure IIa: Log of Total Votes Cast in NLRB Election, by Vote Share**



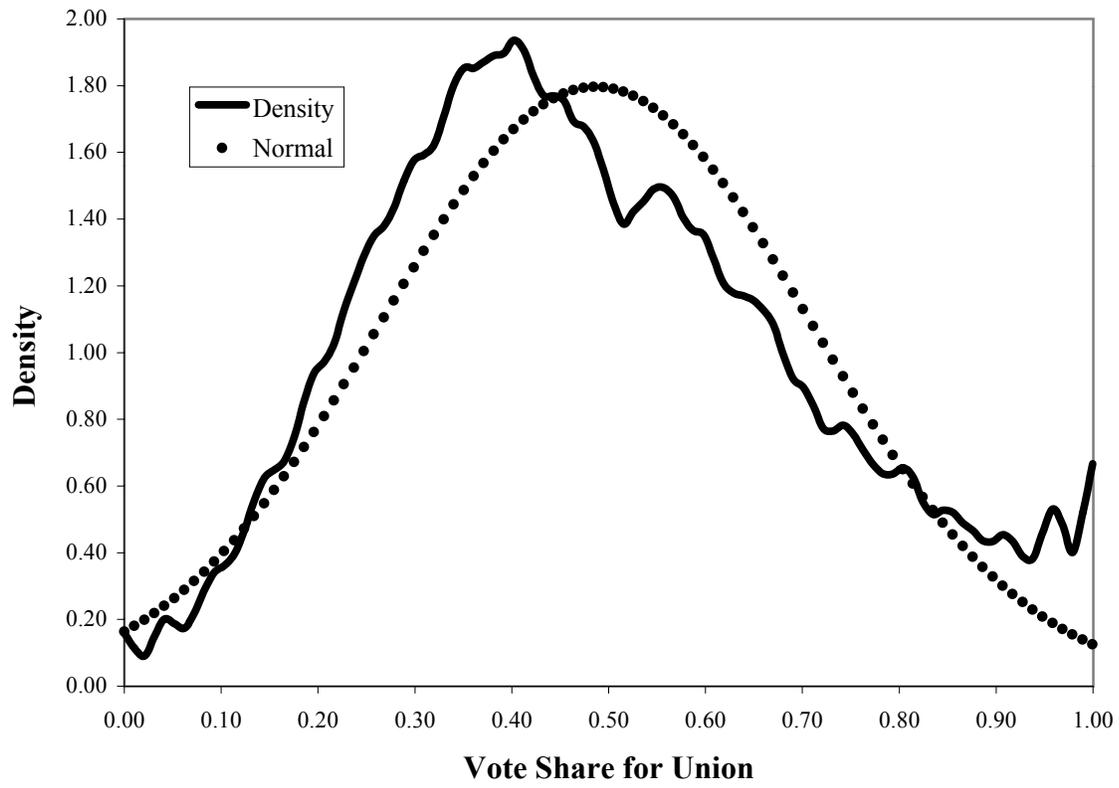
**Figure IIb: Probability in Manufacturing (before Election), by Vote Share**



**Figure IIc: Probability in Service Sector (before Election), by Vote Share**



**Appendix Figure I: Estimated Density of Union Vote Share**



**Table I: Establishment Outcomes (as of 2001) by Election Outcome: NLRB Elections 1983-1999**

Year of Election	Proportion Survived as of 2001				Mean Employment as of 2001				Mean Log of Employment as of 2001			
	All	Loss	Win	Diff.	All	Loss	Win	Diff.	All	Loss	Win	Diff.
<=1983	0.304	0.318	0.280	-0.037	65.4	80.4	40.8	-39.5	4.35	4.55	3.99	-0.55
1984	0.277	0.294	0.255	-0.039	61.4	68.8	51.0	-17.8	4.38	4.48	4.22	-0.26
1985	0.312	0.336	0.276	-0.060	57.6	67.3	43.2	-24.1	4.35	4.39	4.27	-0.12
1986	0.309	0.333	0.275	-0.058	53.2	61.6	41.2	-20.4	4.33	4.39	4.21	-0.19
1987	0.327	0.348	0.302	-0.045	57.4	67.1	45.3	-21.9	4.39	4.51	4.22	-0.29
1988	0.360	0.372	0.343	-0.029	68.6	73.6	61.9	-11.7	4.43	4.56	4.25	-0.31
1989	0.372	0.380	0.361	-0.019	75.8	76.7	74.7	-1.9	4.40	4.52	4.24	-0.28
1990	0.393	0.395	0.390	-0.005	66.2	70.3	60.4	-9.9	4.25	4.39	4.06	-0.33
1991	0.411	0.429	0.385	-0.044	84.9	93.7	71.5	-22.3	4.37	4.46	4.23	-0.23
1992	0.432	0.450	0.409	-0.040	80.0	84.7	73.7	-11.0	4.40	4.45	4.32	-0.13
1993	0.416	0.429	0.401	-0.027	75.1	78.0	71.4	-6.5	4.37	4.50	4.18	-0.32
1994	0.465	0.458	0.474	0.016	105.4	102.5	109.3	6.7	4.53	4.61	4.43	-0.18
1995	0.509	0.521	0.491	-0.030	107.1	103.6	112.0	8.4	4.59	4.60	4.57	-0.02
1996	0.522	0.525	0.519	-0.007	107.5	105.5	110.0	4.5	4.47	4.48	4.45	-0.03
1997	0.572	0.613	0.521	-0.092	133.7	153.8	108.9	-44.8	4.57	4.67	4.41	-0.26
1998	0.572	0.578	0.565	-0.013	119.8	111.6	129.9	18.2	4.43	4.52	4.31	-0.21
1999	0.578	0.610	0.542	-0.068	115.5	134.0	94.7	-39.3	4.42	4.56	4.24	-0.32
All	0.417	0.430	0.400	-0.030	83.4	88.3	76.8	-11.5	4.42	4.51	4.30	-0.22
	(0.003)	(0.004)	(0.005)	(0.006)	(1.7)	(2.2)	(2.7)	(3.5)	(0.01)	(0.02)	(0.02)	(0.03)
Obs.	27622	15818	11804		26355	15104	11251		10265	6092	4173	

Note: Standard errors in parentheses. Entries are outcomes, as of 2001, of establishments that experienced an NLRB certification election in a given year. In the middle four columns, zero is assigned to the employment of "dead" establishments; 27622-26355=1267 establishments survived but did not report employment levels. "Loss" and "Win" indicates the outcome of the NLRB representation election. Details of the sample selection criteria are in the Data Appendix.

**Table II: Means of Establishment and Election Outcomes and Characteristics, by Representation Election Outcome, 1983-1999**

	N	Full Sample	Union Loss	Union Win	Difference
1 Survival (Indicator Variable), 2001	27622	0.417 (0.003)	0.430 (0.004)	0.400 (0.005)	-0.030 (0.006)
2 Employment, 2001	26355	83.4 (1.7)	88.3 (2.2)	76.8 (2.7)	-11.5 (3.5)
3 Log of Employment, 2001	10265	4.42 (0.01)	4.51 (0.02)	4.30 (0.02)	-0.22 (0.03)
4 Sales Volume, 2001	25719	14225 (321)	16250 (454)	11501 (441)	-4750 (633)
5 Log of Sales Volume, 2001	9629	9.34 (0.02)	9.48 (0.02)	9.14 (0.03)	-0.35 (0.04)
6 Presence of Union Post-Election (Indicator Variable)	11532	0.206 (0.004)	0.097 (0.004)	0.363 (0.007)	0.266 (0.005)
7 Presence of Union Pre-Election (Indicator Variable)	11532	0.129 (0.003)	0.095 (0.004)	0.179 (0.006)	0.084 (0.005)
8 Number of Eligible Voters	27622	104.1 (0.8)	113.4 (1.2)	91.6 (1.1)	-21.8 (1.6)
9 Log of Eligible Voters	27622	4.22 (0.01)	4.29 (0.01)	4.14 (0.01)	-0.15 (0.01)
10 Number of Votes Cast	27622	91.7 (0.7)	101.9 (1.0)	78.0 (0.9)	-23.9 (1.4)
11 Log of Votes Cast	27622	4.10 (0.01)	4.18 (0.01)	3.99 (0.01)	-0.19 (0.01)
12 Manufacturing Sector (Indicator Variable)	27622	0.380 (0.003)	0.421 (0.004)	0.326 (0.004)	-0.094 (0.006)
13 Service Sector (Indicator Variable)	27622	0.273 (0.003)	0.218 (0.003)	0.348 (0.004)	0.130 (0.005)
14 Trucking Voting Unit (Indicator Variable)	27622	0.150 (0.002)	0.174 (0.003)	0.119 (0.003)	-0.055 (0.004)
15 Log of State Employment, Election Year	27622	15.08 (0.01)	15.06 (0.01)	15.10 (0.00)	0.04 (0.01)
16 Log of State Employment, 2000	27622	15.19 (0.01)	15.18 (0.01)	15.21 (0.02)	0.04 (0.02)
17 Change In Log Emp. (2000 - Election Year)	27622	0.115 (0.001)	0.117 (0.001)	0.113 (0.008)	-0.005 (0.008)
18 State Unemployment Rate, Election Year	27622	6.23 (0.01)	6.24 (0.02)	6.22 (0.02)	-0.02 (0.02)
19 State Unemployment Rate, 2000	27622	4.14 (0.01)	4.11 (0.01)	4.17 (0.01)	0.06 (0.01)
20 Change in UR (2000 - Election Year)	27622	-2.09 (0.01)	-2.13 (0.01)	-2.04 (0.02)	0.08 (0.02)

Note: Standard errors in parentheses. Details of the merged data from the NLRB, FMCS, and InfoUSA are in the Data Appendix. Rows 2 and 4 impute 0 for non-surviving establishments (some data are missing for surviving establishments). Rows 6 and 7 contain only the establishments that survive to the year 2001. Presence of Union post-election (pre-election) indicates whether or not a union at the location of the establishment filed a contract expiration between the election date and 2001 (between the beginning of the FMCS data and the date of the election).

**Table III: Establishment Survival, Union Presence, and Pre-determined Characteristics, by Vote Share for the Union**

	Full Sample			0.25<= Vote Share <= 0.75			0.35<= Vote Share <= 0.65			0.45<= Vote Share <= 0.55		
	Loss	Won	Diff.	Loss	Won	Diff.	Loss	Won	Diff.	Loss	Won	Diff.
Survival, 2001	0.430 (0.004)	0.400 (0.005)	-0.030 (0.006)	0.436 (0.005)	0.413 (0.005)	-0.023 (0.007)	0.438 (0.006)	0.424 (0.007)	-0.014 (0.009)	0.442 (0.011)	0.435 (0.011)	-0.007 (0.015)
Eligible Voters	113 (1)	92 (1)	-22 (2)	120 (1)	99 (1)	-21 (2)	118 (2)	106 (2)	-12 (3)	116 (3)	114 (3)	-2 (5)
Log(Elig. Voters)	4.29 (0.01)	4.14 (0.01)	-0.15 (0.01)	4.34 (0.01)	4.20 (0.01)	-0.14 (0.01)	4.33 (0.01)	4.25 (0.01)	-0.08 (0.02)	4.31 (0.02)	4.32 (0.02)	0.00 (0.03)
Votes Cast	102 (1)	78 (1)	-24 (1)	108 (1)	87 (1)	-21 (2)	107 (2)	93 (2)	-13 (2)	105 (3)	102 (3)	-3 (4)
Log(Votes Cast)	4.18 (0.01)	3.99 (0.01)	-0.19 (0.01)	4.24 (0.01)	4.08 (0.01)	-0.16 (0.01)	4.23 (0.01)	4.13 (0.01)	-0.10 (0.01)	4.22 (0.02)	4.22 (0.02)	0.00 (0.03)
Manufacturing	0.421 (0.004)	0.326 (0.004)	-0.094 (0.006)	0.437 (0.005)	0.363 (0.005)	-0.074 (0.007)	0.436 (0.006)	0.372 (0.006)	-0.064 (0.009)	0.422 (0.010)	0.397 (0.011)	-0.025 (0.015)
Service Sector	0.218 (0.003)	0.348 (0.004)	0.130 (0.005)	0.223 (0.004)	0.322 (0.005)	0.099 (0.006)	0.234 (0.005)	0.315 (0.006)	0.081 (0.008)	0.251 (0.009)	0.290 (0.010)	0.038 (0.014)
Trucking	0.174 (0.003)	0.119 (0.003)	-0.055 (0.004)	0.162 (0.003)	0.128 (0.004)	-0.034 (0.005)	0.148 (0.004)	0.125 (0.004)	-0.023 (0.006)	0.135 (0.007)	0.124 (0.008)	-0.011 (0.010)
Number of Obs.	15818	11804		11851	8112		7351	5557		2217	1923	
Union Present	0.097 (0.004)	0.363 (0.007)	0.266 (0.007)	0.105 (0.004)	0.344 (0.008)	0.239 (0.008)	0.117 (0.006)	0.332 (0.010)	0.216 (0.011)	0.117 (0.010)	0.276 (0.015)	0.159 (0.018)
Post-Election	0.095 (0.004)	0.179 (0.006)	0.084 (0.006)	0.097 (0.004)	0.155 (0.006)	0.058 (0.007)	0.101 (0.005)	0.149 (0.007)	0.048 (0.009)	0.099 (0.010)	0.132 (0.012)	0.033 (0.015)
Pre-Election												
Number of Obs.	6806	4726		5163	3350		3222	2358		980	836	

Note: Standard errors in parentheses. Entries are means and differences by election outcome, in the full and, three separate sub-samples, by election vote share. Upper panel includes establishments, whether or not they survive as of the year 2001, and the lower panel includes only those establishments that survive as of the year 2001. "Union Present post-election" (pre-election) indicates whether or not a union at the location of the establishment filed a contract expiration between the election date and 2001 (between the beginning of the FMCS data and the date of the election). Details of data set construction in the Data Appendix.

**Table IV: Reduced Form Specification: Effect of Union Victory on Probability of Survival (2001), and Post-Election Presence of Union**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Probability of Survival							
Union Victory	-0.007 (0.014)	-0.006 (0.014)	-0.003 (0.014)	-0.004 (0.013)	-0.008 (0.014)	-0.008 (0.013)	-0.007 (0.013)
Presence of Union (Pre-Election)	---	---	---	---	0.094 (0.010)	0.093 (0.010)	---
Log(Eligible Vote)	---	---	---	---	0.022 (0.004)	0.012 (0.004)	---
Year Dummies	No	Yes	Yes	Yes	Yes	Yes	---
State Dummies	No	No	Yes	Yes	Yes	Yes	---
Industry Dummies	No	No	No	Yes	No	Yes	---
Unit Dummies	No	No	No	Yes	No	Yes	---
Probability of Post-Election Union Present							
Union Victory	0.152 (0.017)	0.150 (0.017)	0.143 (0.017)	0.145 (0.017)	0.131 (0.016)	0.135 (0.016)	0.132 (0.016)
Presence of Union (Pre-Election)	---	---	---	---	0.401 (0.013)	0.394 (0.013)	---
Log(Eligible Vote)	---	---	---	---	0.015 (0.004)	0.006 (0.004)	---
Year Dummies	No	Yes	Yes	Yes	Yes	Yes	---
State Dummies	No	No	Yes	Yes	Yes	Yes	---
Industry Dummies	No	No	No	Yes	No	Yes	---
Unit Dummies	No	No	No	Yes	No	Yes	---

Note: N=27622 for upper panel, N=11532 (includes only those establishments that survive by 2001) for lower panel. Robust standard errors in parentheses. Upper panel refers to survival, as of 2001, as the dependent variable; lower panel refers to the observation of a contract expiration after the election. All specifications are least squares regressions, for the upper (lower) panel include a 4th order polynomial in the vote share, a dummy variable for win/loss and the dummy variable interacted with the linear term in the vote share. Column (7) regresses the residuals - from an initial regression of survival (presence of union) on all the covariates - on a 4th order polynomial in the vote share, a dummy variable for win/loss and the dummy variable interacted with the linear term in the vote share. Details of the data set in the Data Appendix.

**Table V: Reduced-Form Results: Impact of Certification on Establishment Outcomes, Overall, and by Year of Election, Industry, and Voting Unit Size**

Dependent Variable	Survival	Survival (2)	Empl.	Log(Empl.)	Sales	Log(Sales)
Mean	0.417	0.643	83	4.42	14225	9.34
(Std. Dev)	(0.493)	(0.479)	-276	-1.46	-51542	-1.67
Overall Effect	-0.008	0.012	0	-0.07	250	-0.03
	(0.013)	(0.013)	(8)	(0.06)	(1470)	(0.07)
Year: before 1988 (Main)	-0.022	0.009	-17	-0.18	-326	-0.10
	(0.020)	(0.021)	(10)	(0.11)	(1889)	(0.12)
Year: 1988-1991 (Interaction)	0.001	0.003	13	0.17	-777	0.10
	(0.028)	(0.028)	(14)	(0.15)	(2708)	(0.16)
Year: 1992-1995 (Interaction)	0.034	0.013	26	-0.02	2065	-0.02
	(0.029)	(0.029)	(16)	(0.14)	(3013)	(0.16)
Year: after 1995 (Interaction)	0.022	0.003	34	0.25	1536	0.17
	(0.029)	(0.028)	(17)	(0.14)	(3349)	(0.15)
Other Industry (Main)	-0.009	0.018	-2	0.06	1641	0.06
	(0.021)	(0.020)	(9)	(0.09)	(2087)	(0.12)
Manufacturing (Interaction)	0.002	-0.013	2	-0.24	-3887	-0.24
	(0.025)	(0.024)	(12)	(0.10)	(2857)	(0.13)
Service (Interaction)	-0.016	-0.004	3	-0.07	917	0.04
	(0.027)	(0.026)	(17)	(0.14)	(2551)	(0.16)
El. Vote: < 40	0.005	0.023	-5	-0.11	2308	0.00
	(0.022)	(0.021)	(10)	(0.09)	(1861)	(0.12)
El. Vote: 40-100	-0.032	-0.005	-4	0.02	-1223	-0.03
	(0.026)	(0.025)	(11)	(0.11)	(2040)	(0.13)
El. Vote: > 100	-0.008	-0.028	30	0.06	-4318	-0.09
	(0.028)	(0.027)	(18)	(0.13)	(3476)	(0.16)
Number of Obs.	27622	27622	26355	10265	25719	9629

Note: Standard Deviations in first row, robust standard errors otherwise. Survival(2) denotes whether any establishment was present at the exact street address as of 2001, third and fifth columns assign "0" to the "dead" establishments. Estimated Annual Sales Volume is in thousands of dollars. Specifications: Base specification is Column (6) of Table IV. See text for details on interaction specifications. See Data Appendix for details on data set construction

**Appendix Table I: Summary Statistics for Merged NLRB-FMCS-InfoUSA data, full sample and sample for estimation**

Variable	All Certification Elections			Sample for Estimation		
	Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Survival (Indicator Variable), 2001	51132	0.394	0.489	27622	0.417	0.493
Survival (2) (Indicator Variable), 2001	51132	0.623	0.485	27622	0.643	0.479
Employment, 2001	48836	66	290	26355	83	276
Log of Employment, 2001	17863	4.02	1.54	10265	4.42	1.46
Sales Volume, 2001	29367	28603	103667	16349	32865	74884
Log of Sales Volume, 2001	29367	8.68	1.83	16349	8.99	1.84
Union Present, Post-Election	51132	0.184	0.388	27622	0.175	0.380
Union Present, Pre-Election	51132	0.117	0.321	27622	0.102	0.303
Eligible Voters	51132	63	111	27622	104	135
Log(Elig. Voters)	51132	3.33	1.25	27622	4.22	0.83
Votes Cast	51132	55	98	27622	92	119
Log(Votes Cast)	51083	3.20	1.25	27622	4.10	0.82
Manufacturing	51132	0.306	0.461	27622	0.380	0.485
Service Sector	51132	0.260	0.439	27622	0.273	0.446
Trucking	51132	0.177	0.382	27622	0.150	0.357
Log(State Emp.), Year of Election	50284	15.08	0.87	27622	15.08	0.85
Log(State Emp.), Year 2000	50284	15.19	0.87	27622	15.19	0.85
Change in Log(Emp.)	50284	0.115	0.097	27622	0.115	0.096
State UR, Year of Election	50284	6.2	1.8	27622	6.2	1.8
State UR, Year 2000	50284	4.2	0.9	27622	4.1	0.8
Change in UR	50284	-2.1	1.7	27622	-2.1	1.7

Note: Sample restriction is that the number of votes cast be greater than or equal to 20, and the establishment is within the 50 U.S. states and District of Columbia. Details of data set construction in Data Appendix

**Appendix Table II: NLRB Elections: Win rates and Average Union Vote Share, 1983 - 1999**

---

Year of Election	Obs.	Winning Percentage	Vote Share for Union	
			Mean	Std Dev.
<=1983	721	0.376	0.463	0.199
1984	1802	0.416	0.474	0.220
1985	1810	0.403	0.471	0.217
1986	1793	0.408	0.476	0.222
1987	1768	0.445	0.494	0.229
1988	1809	0.425	0.493	0.227
1989	1890	0.442	0.496	0.229
1990	1846	0.413	0.488	0.224
1991	1648	0.397	0.474	0.221
1992	1446	0.431	0.488	0.223
1993	1710	0.447	0.499	0.222
1994	1625	0.431	0.485	0.226
1995	1508	0.417	0.477	0.215
1996	1629	0.427	0.487	0.218
1997	1760	0.445	0.492	0.222
1998	1741	0.455	0.505	0.218
1999	1116	0.470	0.511	0.226
All	27622	0.427	0.487	0.222

---

Note: Restricted sample of elections with 20 or more valid votes cast. See Appendix Table I.

**Appendix Table III: Union Presence, Pre-determined Characteristics for Surviving Establishments, by Vote Share for the Union**

	Full Sample			0.25<= Vote Share <= 0.75			0.35<= Vote Share <= 0.65			0.45<= Vote Share <= 0.55		
	Loss	Won	Diff.	Loss	Won	Diff.	Loss	Won	Diff.	Loss	Won	Diff.
Union Present	0.097	0.371	0.274	0.103	0.350	0.247	0.116	0.338	0.222	0.120	0.275	0.155
Post-Election	(0.004)	(0.007)	(0.008)	(0.004)	(0.009)	(0.009)	(0.006)	(0.010)	(0.011)	(0.011)	(0.016)	(0.019)
Union Present	0.092	0.180	0.087	0.095	0.152	0.058	0.101	0.145	0.044	0.102	0.126	0.024
Pre-Election	(0.004)	(0.006)	(0.007)	(0.004)	(0.007)	(0.008)	(0.006)	(0.008)	(0.009)	(0.010)	(0.012)	(0.016)
Eligible Voters	120	99	-21	125	107	-18	124	114	-10	123	126	3
	(2)	(2)	(3)	(2)	(3)	(4)	(3)	(3)	(5)	(6)	(6)	(8)
Log(Elig. Voters)	4.33	4.18	-0.15	4.37	4.24	-0.13	4.35	4.29	-0.07	4.35	4.38	0.03
	(0.011)	(0.013)	(0.017)	(0.013)	(0.015)	(0.020)	(0.017)	(0.019)	(0.025)	(0.029)	(0.032)	(0.044)
Votes Cast	108	85	-23	113	94	-19	112	101	-11	110	113	3
	(2)	(2)	(3)	(2)	(2)	(3)	(3)	(3)	(4)	(5)	(5)	(7)
Log(Votes Cast)	4.22	4.04	-0.19	4.27	4.13	-0.15	4.26	4.18	-0.08	4.25	4.28	0.03
	(0.011)	(0.012)	(0.017)	(0.013)	(0.015)	(0.020)	(0.016)	(0.018)	(0.025)	(0.029)	(0.032)	(0.043)
Manufacturing	0.450	0.335	-0.115	0.464	0.373	-0.091	0.456	0.380	-0.075	0.420	0.403	-0.017
	(0.006)	(0.007)	(0.010)	(0.007)	(0.009)	(0.012)	(0.009)	(0.011)	(0.014)	(0.017)	(0.018)	(0.025)
Service Sector	0.234	0.380	0.147	0.243	0.352	0.109	0.259	0.339	0.079	0.277	0.310	0.033
	(0.005)	(0.008)	(0.009)	(0.006)	(0.009)	(0.011)	(0.008)	(0.010)	(0.013)	(0.015)	(0.017)	(0.023)
Trucking	0.168	0.112	-0.055	0.157	0.123	-0.034	0.146	0.120	-0.025	0.131	0.118	-0.013
	(0.005)	(0.005)	(0.007)	(0.005)	(0.006)	(0.008)	(0.007)	(0.007)	(0.010)	(0.011)	(0.012)	(0.016)
Number of Obs.	6092	4173		4630	2949		2876	2079		886	739	

Note: Standard errors in parentheses. Entries are means and differences by election outcome, in the full and, three separate sub-samples, by election vote share. Observations include only those establishments that survive as of the year 2001, and have nonmissing employment. "Union Present post-election" (pre-election) indicates whether or not a union at the location of the establishment filed a contract expiration between the election date and 2001 (between the beginning of the FMCS data and the date of the election). Details of data set construction in the Data Appendix.