NBER WORKING PAPER SERIES

UNION SUCCESS IN REPRESENTATION ELECTIONS: WHY DOES UNIT SIZE MATTER?

Henry S. Farber

Working Paper 7229
http://www.nber.org/papers/w7229

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 1999

Union Success in Representation Elections:
Why Does Unit Size Matter?
Henry S. Farber
NBER Working Paper No. 7229
July 1999
JEL No. J50, J51

ABSTRACT

I establish four facts regarding the pattern of NLRB supervised representation election activity over the past 45 years: 1) the quantity of election activity has fallen sharply and discontinuously since the mid-70's after increasing between the mid-1950's and the mid-1970's; 2) union success in elections held has declined less sharply, though continuously, over the entire period; 3) it has always been the case that unions have been less likely to win NLRB-supervised representation elections in large units than in small units; and 4) the size-gap in union success rates has widened substantially over the last forty years. I develop a simple optimizing model of the union decision to hold a representation election that can account for the first three facts. I provide a pair of competing explanations for the fourth fact: one based on differential behavior by employers of different sizes and one purely statistical. I then develop and estimate three empirical models of election outcomes using data on NLRB elections over the 1952-98 time period in order to determine whether the simple statistical model can account for the size pattern of union win rates over time. I conclude that systematic union selection of targets for organization combined with the purely statistical factors can largely account for observed patterns.

Henry S. Farber
Industrial Relations Section
Firestone Library
Princeton University
Princeton, NJ 08544
and NBER
farber@princeton.edu

Union Success in Representation Elections: Why Does Unit Size Matter?

Henry S. Farber

1. Introduction

The decline in unionization in the United States in the last twenty-five years has been accompanied by a concomitant decline in amount of union organizing activity as measured by representation elections administered by the National Labor Relations Board (NLRB).[1] Over the longer term, election activity increased sharply during the twenty years from 1955-1975 and then declined sharply over the seven years between about 1976 and 1983. In contrast, union success in the representation elections that have been held has declined continuously since the late 1940's.[2] One striking feature of the decline in union success in winning these elections is that, while unions have always been less likely to win large elections than small elections, the union success rate fell substantially more in large units than in small units. The task I undertake here is to provide a theoretical framework to understand these facts and to investigate competing explanations for the differential decline in win rates by size category.

I establish these facts in the next two sections using data on aggregate trends in election activity and outcomes in the post-World War II period. In section 4, I develop a simple optimizing model of the union decision to hold a representation election. This model has two implications. First, in response to changes in the economic, political, and social

Prepared for the Cornell-Princeton Eighth Annual Policy Conference, "New Research Perspectives on Labor Unions," May 1999. Useful comments were received from participants that this conference and from participants in workshops at Princeton, the National Bureau of Economic Research, and a conference celebrating the 50th anniversary of the Industrial Relations Center of the University of Hawaii. Gena Estes, Bridget Terry, and Abigail Payne provided able assistance in compiling the data used in this study.

[1] There is a substantial literature measuring and attempting to account for the decline in unionization. See, for example, Freeman and Medoff (1984), Dickens and Leonard (1985), Farber (1985), Kochan, Katz, and McKersie (1986), Farber (1990), Riddell (1993), Farber and Krueger (1993), and Western (1997).

[2] Weiler (1983) and Freeman (1985) emphasize the role that illegal employer resistance to unions prior and during election campaigns has played in the decline of union success in representation elections. LaLonde and Meltzer (1991) argue that employer illegalities have not been an important factor in elections, and Weiler (1991) responds. Dickens (1983) presents direct evidence that illegal employer behavior during election campaigns can have a significant effect on election outcomes.

environment that affect the likelihood of a union election win, unions will adjust the quantity of election activity. But union success in the elections actually held will not be greatly affected. This can reconcile the very different time-series of the quantity of election activity and the union win rate. Second, unions will have lower win rates in large elections than in small elections.

In section 5, I present two competing explanations for the divergence over time in win rates between large units and small units. A "behavioral" explanation for the difference in win rates is that employers in larger establishments have become relatively more effective at mounting anti-union campaigns when an election is scheduled so that there is relatively less sentiment for unions on election day. An alternative and more parsimonious explanation relies on changes over time in the probability of voting for union representation in elections. Suppose (as is, in fact, the case) that over time the fraction of voters in NLRB elections casting pro-union votes has fallen from greater than 0.5 on average to less than 0.5 on average. Suppose further that this decline is of the same magnitude in both large and small units (in contrast to what would happen if large employers are waging more effective campaigns). It is straightforward to show on a purely probabilistic basis that the union win rate will fall more in large elections than in small elections.

Section 6 contains an analysis of three statistical models of union win rates by size category that relies on aggregate election data by size category over the 1952-1998 period. I use this analysis to determine if the differential decline in union win rates by unit size can be accounted for without allowing for a differential decline in the distribution of pro-union vote shares by unit size. A finding that the differential decline in win rates can, in fact, be accounted for in this manner is evidence in support of the parsimonious statistical explanation. I find that a model constraining the mean difference in the pro-union vote probability between large and small units to be fixed over time but allowing for common time effects and random heterogeneity across elections in the pro-union vote probability accounts for the time-series patterns of win rates quite well.

The final section contains some concluding observations on the analysis.

2. Overall Trends in Election Activity and Union Success

In order to examine overall trends in election activity and outcomes, I use the tabular information from the NLRB annual reports since 1946 in combination with comparable statistics from the micro-data on individual elections from 1973 to 1998.[3] I present some simple statistics on the time-series on election activity and union win rates over time by size of unit.

Figure 1 documents the dramatic decline in the number of representation elections held during the post-war period (1946-1998). Although there is fairly wide year-to-year variation, the time-series of the number of elections shows no obvious trend between 1946 and 1960. After 1960 the number of elections per year increased steadily from 6380 in 1960 to a peak of 8799 in 1973. The number of election then declined slowly through 1980 at which point there was a precipitous decline of over 50 percent in the number of elections between 1980 and 1983 (from 7496 in 1980 to 3478 in 1983). After 1983, the number of elections declined slowly to about 3000 in 1995. The same pattern of long-run decline in election activity is also evident in figure 2, which is a time-series of the total number of votes cast in representation elections.

These trends are even more striking given the facts that private-sector non-agricultural employment in the United States grew from 63 million workers to 103 million workers between 1973 and 1997 while the fraction of the private sector civilian workforce that was unionized fell over the same period from 23.7 percent 10.6 percent.[4] Using these figures, the size of the nonunionized private-sector workforce grew from 48.1 million workers in 1973 to 92.1 million in 1997. The data underlying figure 2 show that 454,358 workers voted in NLRB-supervised representation elections in 1973 while only 186,532 workers voted in such elections in 1997. Thus, fraction of the private-sector nonunionized workforce voting in representation elections fell from 0.9 percent in 1973 to 0.2 percent in 1997.

[3] The dates refer to fiscal years. For consistency over time, fiscal years are defined throughout as running from July through June despite the fact that the federal government switched to an October through September fiscal beginning with the 1977 fiscal year. I use information from the NLRB annual reports rather than from the micro-data for fiscal 1974 because, for unknown reasons, the micro-data provided by the NLRB have no information on elections closed in July, August, and September 1973.

[4] These numbers are based on tabulations of the May 1973 Current Population Survey and the Merged Outgoing Rotation Group data from the 1997 Current Population Survey.

In contrast to the inverted U shape of the time series on election activity, the size of the elections contested fell steadily over time through about 1970, after which average size remained relatively stable. This is documented in figure 3, which shows that the average number of votes cast per election fell from an average of 116 between 1946 and 1950 to a low of 47 in 1983 before increasing to an average of 56 between 1995 and 1998. One explanation for the declining trend might be that the number of large non-unionized establishments that were reasonable targets for organization fell over time as the union movement grew in the post-war period until the 1960's.

Union success in representation elections held declined steadily over the post-war period until 1975, as documented in figure 4. The union win rate fell from 74.6 percent over the 1946-1950 period to 48.2 percent over the 1976-1980 period. The union win rate has been roughly stable since 1980 at about 48 percent. Figure 4 also contains information on the fraction of votes cast in representation elections each year that were cast in favor of union representation. This time series shows the same decline over the post-war period that is shown by the win rate. In fact, the pro-union vote share and the win rate appear to be virtually identical (correlation 0.96 with almost identical means of 0.597 for the win rate and 0.604 for the pro-union vote share).[5] In contrast to the time series on the quantity of election activity, neither the win rate nor the pro-union vote share show a sharp discontinuity in the early 1980's.

3. Unit Size, Election Activity, and Election Outcomes

In comparing the time series on average election size (figure 3) and win rates (figure 4), it appears that union success fell as election size fell. The correlation of these two time series is 0.705. However, it would be incorrect to infer that the decline in union success over time is due to a change in the mix of elections from large elections to small elections. In fact, it is not the case that unions have been more successful in large elections than in small elections.

[5] The convergence of the union win rate and the pro-union vote share presents a puzzle that I will address in the analysis of the statistical model.

Figure 5 contains union win rates in representation elections for three size categories (1-9 votes, 10-99 votes, ≥100 votes), and the underlying data for five size categories (1-9 votes, 10-19 votes, 20-49 votes, 50-99 votes, ≥100 votes) are presented in table 1. It is uniformly true that union win rates decline with election size. For the purposes of this discussion, I call the difference in win rates between the smallest and largest size categories (1-9 voters and ≥ 100 voters respectively) the size-gap in win rates. The size gap in win rates was relatively small in 1952 (win rates of 0.78 vs. 0.69 for a gap of 0.09), but the size-gap increased steadily to a peak in 1978 and 1979 (win rates of 0.59 vs. 0.31 for a gap of 0.28). The size gap in win rates has been relatively stable since that time. The size gap in win rates tripled in the quarter century between 1952 and 1978, and has not increased since. Another way to view the data is that union win rates declined over this period in all size categories but that the decline was largest in the larger size categories. The union win rate in the smallest category fell by 20 percentage points between 1952 and 1977 (from 0.78 to 0.58), but the win rate fell by 33 percentage points (from 0.69 to 0.36) over the same period in the largest category.

It is natural to ask how the pro-union vote shares by size category match the win rates. Unfortunately, information on the pro-union vote share by size category is available only since 1973 when micro-data on individual elections became available, and about half of the increase in the size-gap in win rates occurred prior to 1973 with the rest of the increase occurred between 1973 and 1978.[6] Figure 6 contains a graph of average pro-union vote shares since 1973 for three size categories (1-9 votes, 10-99 votes, ≥100 votes), and the underlying data for five size categories ((1-9 votes, 10-19 votes, 20-49 votes, 50-99 votes, ≥100 votes) are presented in table 2.[7] The size gap in the pro-union vote share was 0.12 in 1973 and increased slightly to 0.14 1978, when the size gap in the win rate was largest. Over the next five years, the size gap in the vote share fell sharply to 7 percentage points

[6] Reading from table 1, the size gap in win rates was 0.09 in 1952, 0.19 in 1973, 0.24 in 1974, 0.25 in 1976, and 0.28 in 1978, and 0.25 in 1995.

[7] Note that what are plotted in figure 6 and presented in table 2 are *not* the ratios of the number of total pro-union votes cast to the total number of votes cast. The figure and table contain the *average* across elections of the ratio of pro-union votes to total votes. There are no data for fiscal 1974 because the data file provided by the NLRB for that year is missing data on all elections closed in July, August, and September 1973.

in 1983 while the size gap in the union win rate fell only slightly. Subsequently, the the size gap in the pro-union vote share increased to 15 percentage points in 1998 with a 9 point increase in the size-gap in the union win rate.

I noted when discussing figure 4 that the share of votes cast in favor of union representation tracked the union win rate very closely over the post-war period. I present graphs by size category of the union win rate and the average union vote share in figure 7. The upper left hand panel is for all elections, and it shows, like figure 4, that the win rate and vote shares track quite closely.[8] The remaining three panels of figure 7 are specific to the three size categories, and they show that the win rates track the vote shares more closely in small elections than in large elections. In large elections, the average pro-union vote share is close to 0.5 while the union win rate diverges sharply from 0.5.[9]

4. An Economic Model of the Decision to Hold Representation Elections

In this section, I present an economic model of the union decision to hold a representation election that can account for several of the facts established in the previous two sections.

A rational labor union will contest elections only where there is a positive expected value associated with the election. This suggests that among all possible potential bargaining units, called "targets" here, elections are more likely to be held where the likelihood of a union victory is higher. This has some important implications for the analysis of both the quantity of election activity and election outcomes over time. First, the elections observed at any point in time are not representative of the pool of targets. Second, the pool of "favorable" targets will become depleted over time as unions win elections in the best targets. This is true 1) even if the likelihood of union victory within particular targets is not changing over time and 2) new targets are being created due to the creation of new firms and the growth of existing firms into new establishments. Thus, independent of

[8] The data on vote shares in figure 4 are the ratios of total prounion votes cast in all elections to total votes cast in all elections. The data on vote shares in figure 7 are the averages across elections of the election specific pro-union vote shares. The latter require the election level micro-data to compute and are only available since 1973.

[9] These general patterns are consistent with a simple binomial model of voting of the sort I discuss below.

any changes in worker or employer attitudes, the quantity of union organizing is likely to decline over time.

I characterize each target in two dimensions. The first is the probability that the union will win an election should one be held. Denote this probability by θ_i, where i indexes the target, and it represents the probability that more than one-half of the voters cast their votes in favor of union representation.[10] The second relevant characteristic of the target is its size. The size may have a direct effect on the probability of a union victory. For example, the binomial model I derive below, the probability of a union victory is increasing in size if the probability of an individual pro-union vote is greater than 0.5 and decreasing in size if the probability of an individual pro-union vote is less than 0.5. The number of workers could also have an important bearing on the attractiveness of the target to the union in other ways. A union victory in a large election could have important positive spillovers for the union in terms of bargaining leverage and "marketing" value in other organizing campaigns. Additionally, there may be decreasing costs per worker of holding the organizing drive and/or decreasing costs per member of servicing a bargaining unit once there is a union victory.

Define the expected value to the union of contesting an election at target i as

$$V_i = \theta_i(R_i - C_{ai}) - C_{oi}, \tag{1}$$

where θ_i is the probability that the union wins the election, R_i is the value to the union once target i is organized, C_{ai} is the administrative cost to the union of servicing target i once it is organized, and C_{oi} is the cost of the organizing effort. A rational union will undertake to organize the target if V_i is positive. This implies that the condition for an election to be held is

$$\theta_i > \frac{C_{oi}}{(R_i - C_{ai})}. \tag{2}$$

[10] There is a status-quo bias built into the election process through the National Labor Relations Acts. Thus, unions lose ties in certification elections and unions win ties in decertification elections. Only certification elections are considered here.

The right hand size of equation 5 defines a critical value for the probability of a union victory. This is

$$\theta_i^* = \frac{C_{oi}}{(R_i - C_{ai})}, \qquad (3)$$

and unions will contest elections where $\theta_i > \theta_i^*$.

Now consider the effect of a change in the economic, political, or social environment that changes the distribution of θ without changing θ^*. Suppose there was a change that was adverse to unions so that the distribution of θ shifted to the left; i.e, that there were fewer good targets for organization. The first-order result will be that fewer elections would be held. But, since the selection rule remains unchanged, union success in elections that are held will not be greatly affected[11] Such an adverse change in the distribution of θ could account for the sharp decline in union election activity around 1980 with no corresponding change in union success in elections held.[12]

With regard to the size gap in win rates, on the basis of arguments made above, it is reasonable to assume that the critical value, θ_i^*, is decreasing in size. This can be made clear by dividing both numerator and denominator on the right-hand side of equation 6 by the number of workers in the target (n). This does not change the value of the ratio and the resulting elements can be interpreted as the per-worker value to the union (R_i/n), the per-worker administrative cost (C_{ai}/n), and the per-worker organization cost (C_{oi}/n). The external benefits to the union of organization suggest that R_i/n is increasing in n. The decreasing costs both of holding an organizing drive and of administering a bargaining unit suggest that both C_{oi}/n and C_{ai}/n are declining in n. These are sufficient conditions for θ_i^* to decline with the size of the target. Thus, unions will be willing to undertake elections in larger units where there is a smaller probability of victory.

The model in this section has three implications that are consistent with the facts established in sections 2 and 3. First, win rates decline over time as the the stock of

[11] In fact, the extent to which union success will be affected depends on the underlying distribution of θ before and after the shift.

[12] In fact, there is substantial evidence that political and legal environment for unions worsened substantially in the early 1980's. See Weiler (1990), Gould (1993), and Levy (1985).

high-probability targets is depleted by past organization efforts. Second, changes in the economic, political, and social environment will have sharper effects on the quantity of elections than on election outcomes. Third, there will be an inverse relationship between the union win rate in elections held and unit size.

5. A Statistical Explanation for Divergence in Win Rates by Size

In this section, I develop a simple statistical model that can account for the fact, established in section 3, that the size gap in win rates has increased over time.

Suppose that p_i represents a given worker's probability of voting in favor of union representation. Suppose further, that p_i falls over time by the same amount in elections of different sizes. The binomial model has the clear implication that the effect of this decline on the win rate will differ by the size of election. Assuming a common p and independence of votes across individuals, the number of pro-union votes in an election is distributed as a binomial. The vote share is an unbiased estimate of p in the population, but the win rate will depend strongly on both p and the size of the election. If p is 0.5 then the win rate will also be 0.5, but if p is other than 0.5 then the win rate will diverge from p in predictable ways. If p is greater than 0.5, then the union win rate will be greater than p and the discrepancy between p and win rate will increase with election size.[13] Similarly, if p is less than 0.5, then the union win rate will be less than p and the absolute discrepancy between the win rate and p also increases with election size.

Formally, the probability of there being s pro-union votes in an election with n voters where each voter votes pro-union with probability p is

$$(4) \qquad P(s|n,p) = \binom{n}{s} p^s (1-p)^{n-s}.$$

Since the union wins the election when more than half the votes cast are pro-union, the probability of a union win is

$$(5) \qquad P(win|n,p) = \sum_{s>n/2}^{n} \binom{n}{s} p^s (1-p)^{n-s}.$$

[13] Actually, if there is only one voter in the election, then the win rate is trivially equal to p.

The well-known normal approximation to the binomial, commonly assumed to be valid when $np > 5$ and $n(1-p) > 5$, is that $(s|n,p) \sim N(np, np(1-p))$ so that the probability of a union win is approximately

$$P(win|n,p) = \int_{n/2}^{\infty} \phi\left(\frac{(s-np)}{\sqrt{np(1-p)}}\right) ds \tag{6}$$
$$= 1 - \Phi\left(\frac{\sqrt{n}(1/2-p)}{\sqrt{p(1-p)}}\right),$$

where ϕ and Φ are the standard normal probability density and cumulative distribution functions respectively. It is straightforward to show that $\frac{\partial P(win|n,p)}{\partial n}$ is positive if $p > 1/2$ and negative if $p < 1/2$.

This is a simple application of the law of large numbers applied to the binomial distribution. The expected share of pro-union votes in the binomial case is p for any election size, and the variance of the share is $p(1-p)/n$ where n is the number of voters. Thus the variance of the vote share falls with election size while the mean is unaffected. Consider first the case where p is above 0.5. Since the union wins when the vote share is above 0.5, the probability that the observed vote share is above 0.5 (i.e. not too far below the mean of the distribution) increases with election size because the range of realized vote shares around p narrows with the reduction in variance. Simple calculation based on the binomial with a p of 0.6 shows that the win rate is 0.6 with one voter, 0.74 with 11 voters, and 0.93 with 51 voters. The analogous logic holds when the vote share is less than 0.5. Simple calculation based on the binomial with a p of 0.4 shows that the win rate is 0.4 with one voter, 0.25 with 11 voters, and 0.07 with 51 voters.

The clear implication of this statistical model is that a decline in p could account for the divergence in win rates by size category that was established earlier.

6. Analysis of Win Rates by Size Category - Three Statistical Models

The Saturated Binomial Model

I use the data on win rates in five size categories for 47 years (235 observations), presented in table 1, to estimate a set of simple models of election outcomes. I start by assuming that the probability of a union win is as defined by the binomial model in equation 5. Given data on n_{jt}, the observed probability of a union win is determined by p_{jt}, the probability that a worker in size category j in year t votes for union representation. Each of the 235 observations on the win rate implies a value for $\hat{p}_{jt}$, an estimate of p_{jt}, that can be calculated by solving the implicit relationship in equation 5 for p. Given that the data are aggregated in size categories, specific values for n are required in order to solve for $\hat{p}_{jt}$. I assume that the size distribution of elections within size category that is observed from 1973 through 1998, the period for which I have election-level data, holds for the entire 1952-1998 period. On this basis, I allocate the elections within each size category in each year to specific values of n. I then use this distribution of n and the win rate in size category j in year t together with equation 5 (applied to each specific n within category j) to calculate $\hat{p}_{jt}$. This model is saturated in that there are 235 observations on win rates that imply 235 estimates of the vote probabilities. The model "fits" perfectly in that the 235 win rates are "predicted" perfectly by the 235 vote probabilities.

Figure 8 contains a plot of win rates and predicted vote shares for three size categories using this saturated binomial model as well as a plot of the observed overall win rates and vote shares (similar to figure 7). Only three of the five size categories are presented, and, contrary to my earlier practice, the middle category covers only elections with 20-49 voters. The results for elections with 10-19 voters and 50-99 voters are not presented.[14] The implied vote shares in the smallest election category track the win rates fairly closely. The implied vote shares in the largest category stay very close to 0.5. It takes only slight decline in the vote share from just above 0.5 to just below 0.5 to account for the dramatic

[14] The results for the two categories not presented are as expected in that the results for 10-19 voters lie between those for 1-9 voters and those for 20-49 voters. Similarly, the results for 50-99 voters lie between those for 20-49 voters and those for ≥ 100 voters.

decline in win rates in large elections. These results are a natural consequence of the binomial model.

Figure 9 contains a plot of only the the implied vote shares for the three size categories. It is clear that, while the decline in win rates is steepest for the largest size category, the decline in the vote share necessary to produce this decline is smallest for this category. Thus, there is no evidence using the binomial model that the relatively steep decline in win rates in large elections was due to a relatively steep decline in the prounion vote share. This is not consistent with the explanation for the increased size gap in win rates based on an increase in resistance to union organizing by large employers.

Next, I estimate a constrained binomial model with fewer than 235 parameters in order to provide a more parsimonious description of the data.

The Constrained Binomial Model

I define the constrained binomial model such that p_{jt} has only size category and year main effects. In other words,

$$p_{jt} = \delta_j + \gamma_t, \tag{7}$$

where δ_j is a factor that is specific to size category that does not vary over time and γ_t is a factor that is year-specific and does not vary across size categories. This model has 51 parameters, and it is a constrained version of the saturated model where all of the size-year interactions are assumed to be zero. This is an interesting model because, as is consistent with the null hypothesis that there has not been a steeper decline in vote share in large elections than in small elections, it constrains the differences by size category to be fixed over time. I estimate this model by maximum likelihood, and, while the model is strongly rejected by the data (log L = -1237.3), the results are interesting.[15]

The first column of table 3 contains the estimates of the size effects from this model. Not surprisingly, the estimates imply that the pro-union vote share declines monotonically

[15] That is, the model is rejected against the saturated model which, by definition, has a log-likelihood value of zero. The log-likelihood of -1237.3 in the constrained model implies that the constrained model can be rejected against the saturated model despite the 184 constraints.

with election size. The estimates imply that voters in elections with 10-19 voters are 7.7 percentage points less likely to vote for union representation than are voters in elections with 1-9 voters. This increases to a maximum size gap in the pro-union vote probability of 13.3 percentage points between elections with at least 100 voters and elections with fewer than 10 voters. Note that the measured average size gap between the largest and smallest categories in the pro-union vote share reported in table 2 for the 1973-1998 period is almost identical at 13 percentage points. The upper line in figure 10 represents the estimated year effects from the constrained binomial model (1952=0), and they show a decline in the probability of a prounion vote of about 5 percentage points between 1952 (the base year) and 1982. This was followed by a gradual recovery of about 1 percentage point through 1998.[16]

It is useful to ask how well this model fits the observed win rates. One crude measure is the correlation between the observed win rates and win rates predicted using the estimates of the probability of voting for union representation. This correlation is quite high at 0.958. Figure 11 contains plots of the observed and predicted win rates for three size categories as well as the plot of the observed overall win rate and pro-union vote share. The clearest deviation is in the smallest size category (1-9 voters). The predicted win rate is much flatter and smoother than the observed win rate. This is likely a result of requiring the year effects to be common across size categories. The requirement of fitting the largest size category, where small movements in the probability of a pro-union vote translate into large swings in win rates, no doubt forces there to be relatively little variation over time in the probability of a pro-union vote. This is verified by the fact that the correlation between the observed and predicted win rates is higher in the larger size categories ($\rho = 0.960$ for the smallest size category and increasing to $\rho = 0.989$ in the largest size category).

It is also useful to ask how well the predicted pro-union vote shares match the observed pro-union vote shares over the period for which the latter is observed (1973-1998). Figure 12 contains plots of the observed and predicted vote shares for three size categories, and the results show that the observed vote shares are substantially more variable over time

[16] Standard errors on the estimated time effects in the constrained binomial are in the range of 0.0005 to 0.002.

than are the predicted vote shares, which are constrained to have the same time-series variation across size categories. Additionally, the model generally over-predicts the vote shares in the smallest and largest size categories.

The Beta-Binomial Model

A potentially restrictive assumption underlying the binomial model is that the probability of a pro-union vote is the same in all elections with a size/year cell. Direct evidence that there is heterogeneity in p is implicit in the fact that, with average election sizes of over fifty (figure 3) and vote shares substantially different from 0.5 (figure 4), the win rate and vote share track each other closely. To support win rates bounded away from zero or one in large elections, the value of a fixed p cannot deviate very much from 0.5 regardless of the win rate. Thus, it is surely the case that there is substantial variation in p across potential bargaining units that is likely due to variation in worker interest in union representation. In fact, such variation in p served as the basis of the model I developed of the union's decision regarding whether or not to hold an election in a potential bargaining unit.

Without knowing the distribution of p within size/year categories, any model that accounts for this heterogeneity must rely on arbitrary distributional assumptions. A commonly-used distribution for probabilities is the beta distribution. This distribution has positive density on the unit interval, and it has the additional advantage of yielding a tractable result when mixed with the binomial distribution. When the parameter, p, of the binomial distribution is itself distributed as beta, the resulting distribution is called a beta-binomial, and its density function has a convenient functional form (Evans, Hastings, and Peacock, 1993). Specifically, suppose that the number of pro-union votes, s, in an election with n voters is binomially distributed as in equation 1. Suppose further that the probability that an individual worker votes in favor of union representation (p), which is a parameter of the binomial distribution, is itself distributed as beta such that

$$f(p;a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}p^{a-1}(1-p)^{b-1} \tag{8}$$

where a and b are positive parameters and Γ is the gamma function defined as

$$\Gamma(x) = \int_0^{\infty} exp(-u)u^{x-1}du. \tag{9}$$

The expected value of p is $\mu = \frac{a}{(a+b)}$, and the variance of p is $\sigma^2 = \frac{\mu(1-\mu)}{(a+b+1)}$. On this basis, a convenient reparameterization of the beta distribution has $\mu = \frac{a}{(a+b)}$ and $\alpha = a+b$ where μ is bounded on the unit interval and α is positive. This parameterization yields a density function for p of

$$g(p; \mu, \alpha) = \frac{\Gamma(\alpha)}{\Gamma(\mu\alpha)\Gamma((1-\mu)\alpha)} p^{\mu\alpha-1}(1-p)^{(1-\mu)\alpha-1} \tag{10}$$

where the expected value of p is μ and the variance of p is $\sigma^2 = \frac{\mu(1-\mu)}{(\alpha+1)}$. This beta distribution has flexible functional form. The distribution is unimodal (inverse U-shaped) if $\mu\alpha > 1$ and $(1-\mu)\alpha > 1$. Otherwise, the distribution is bimodal (U- or J- shaped). A special case is that the distribution is uniform if $\alpha = 2$ and $\mu = 0.5$.

Given the beta distribution for p, the beta-binomial distribution of the number of prounion votes is

$$h(s; n, \mu, \alpha) = \binom{n}{s} \frac{\Gamma(\alpha)\Gamma(\mu\alpha + s)\Gamma((1-\mu)\alpha + n - s)}{\Gamma(\mu\alpha)\Gamma((1-\mu)\alpha)\Gamma(n+\alpha)}, \tag{11}$$

and the probability that the union wins the election (received more than half the votes cast) is

$$P(win|n, p) = \sum_{s>n/2}^{n} \binom{n}{s} \frac{\Gamma(\alpha)\Gamma(\mu\alpha + s)\Gamma((1-\mu)\alpha + n - s)}{\Gamma(\mu\alpha)\Gamma((1-\mu)\alpha)\Gamma(n+\alpha)}. \tag{12}$$

The mean pro-union vote share is μ, and the variance of the vote share is $(\frac{\mu(1-\mu)}{n})(\frac{\alpha+n}{\alpha+1})$. It is interesting to compare this with the comparable statistics from the standard binomial distribution where the mean pro-union vote share is p, and the variance of the vote share is $\frac{p(1-p)}{n}$. The parameter μ of the beta distribution is directly analogous to the parameter p of binomial distribution as measuring the expected vote share. The variance of the pro-union vote share in the beta-binomial is larger than the variance in the standard binomial by the multiplicative factor $\frac{\alpha+n}{\alpha+1}$. This is due to the fact that p has a nondegenerate distribution. In the limiting case, where $\alpha \to \infty$, the variance of the beta distribution for p tends to zero and the beta-binomial converges to a binomial distribution with parameter μ. The parameter α controls the degree of "over-dispersion" in the distribution of votes.

In order to estimate the beta-binomial model, I specify the mean of the beta distribution as

$$\mu_{jt} = \delta_j + \gamma_t. \tag{13}$$

This specification is directly analogous to the specification of p as a function of main effects for size category and year that I used in the constrained binomial model in equation 4. I assume that α, the parameter controlling the dispersion of the beta distribution for p, is a constant. This model has 52 parameters (51 main effect estimates and 1 dispersion parameter). Once again, this is a constrained version of the saturated model where all of the size-year interactions are assumed to be zero.[17] I estimate this model by maximum likelihood, and, while the model is once again strongly rejected by the data (log L = -926.8), the beta-binomial model does much better than the constrained binomial model (log L = -1237.3).[18] The constrained binomial model is a special case of the beta-binomial model where α tends to infinity. The maximum-likelihood estimate of α is 15.72 (s.e. 1.36), and, while we cannot test the null hypothesis that $\alpha = \infty$, it is clear that α is significantly less than reasonable positive values (e.g., 20.0).

The second column of table 3 contains the estimates of the size effects from this model. These estimates are very close to those derived from the constrained binomial model. The pro-union vote share declines monotonically with election size. The estimates imply that voters in elections with 10-19 voters are 7.6 percentage points less likely to vote for union representation than are votes in elections with 1-9 voters. This increases to a maximum size gap in the pro-union vote probability of 15.3 percentage points between elections with at least 100 voters and elections with fewer than 10 voters.

The lower line in figure 10 contains the time-series plot of the year effects for the beta-binomial model. In contrast to the upper line, which represents the year effects from

---

[17] Note that it is not possible to estimate a saturated version of the beta-binomial model because that would require 236 parameters (a value of μ for each size-year cell plus a value for α) and there are only 235 observations on win rates. The beta-binomial model is a constrained version of the saturated binomial model discussed earlier.

[18] The beta-binomial model is rejected against the saturated model which, by definition, has a log-likelihood value of zero. The log-likelihood of -926.8 in the beta-binomial model implies that the constrained model can be rejected against the saturated model despite the 183 constraints.

the constrained binomial model, the beta-binomial model shows a substantially larger decline in the mean pro-union vote share between 1952 and 1982 (12.6 percentage points in the beta-binomial compared to 5.1 percentage points for the constrained binomial model. This mean pro-union vote share then recovers by about 3.4 percentage points by 1998.[19] While the year effects for the beta-binomial model decline more steeply than those for the constrained binomial, they are very highly correlated ($\rho = 0.998$).

As with the constrained binomial model, it is useful to ask how well the model fits the observed win rates and pro-union vote shares. The correlation between the observed and predicted win rates is 0.984 (compared with a correlation of 0.959 for the constrained binomial model). Figure 13 contains plots of the observed and predicted win rates for three size categories as well as the plot of the observed overall win rate and pro-union vote share. This model clearly tracks the observed win rates better than does the constrained binomial model (figure 11). The improved fit is most evident in the smallest size category (1-9 voters), where a substantial deviation was evident in figure 11. Figure 14 contains plots, analogous to those in figure 12, of the observed and predicted vote average vote shares (μ from the beta-binomial) for three size categories. The results show that the predicted average vote shares from the beta-binomial model match the observed vote shares much more closely than do the predicted vote shares from the constrained binomial model in figure 12.

How much heterogeneity in p within size-year cells does the beta-binomial model imply? The predicted values of μ in the sample range from 0.437 (the largest size category in 1982) to 0.717 (the smallest size category in 1952). Recall that the variance of p is $\sigma^2 = \frac{\mu(1-\mu)}{(\alpha+1)}$. Using a value of μ of 0.6 for illustration and the estimated value of α of 15.72, the implied standard deviation of p is 0.120, which suggests that there is a substantial amount of variation in p.

The variation in p across elections is illustrated graphically in figure 15, which contains plots of the predicted beta density and distribution functions with $\alpha = 15.72$ for four values of μ (0.4, 0.5, 0.6, and 0.7) that roughly cover the range of predicted μ's in the

[19] Standard errors on the estimated time effects in the beta-binomial model are in the range of 0.002 to 0.006.

sample. The first panel contains the predicted density functions, and the area under each density function to the left of the vertical line at 0.5 represents the fraction of elections held where $p < 0.5$. The density functions move to the left with increasing size as μ falls, and the result is that the fraction of elections held where $p < 0.5$ increases with size.

This relationship is illustrated directly in the second panel of figure 15, which contains the predicted cumulative distribution functions with a vertical line at a pro-union vote probability of 0.5. The horizontal lines are at the intersections of the .5 vote probability line with each of the CDF's so that the fraction of elections held where $p < 0.5$ can be read directly from the vertical axis. Thus, where $\mu = 0.4$, 79 percent of elections are held where $p < 0.5$. This fraction decreases with μ so that where $\mu = 0.7$, only 4.9 percent of elections are held where $p < 0.5$.

I used the estimates of the beta-binomial model in table 3 to calculate the fraction of elections elections held where $p < 0.5$ for each size group and year. This uses the estimates of the size and year effects to generate a specific value of μ_{jt} for each size-year cell. These values of μ_{jt} are then used in conjunction with the estimated value of $\alpha = 15.72$ to compute the probability that a beta distributed random variable with those parameters is less than 0.5. Figure 16 contains plots of this probability for three of the size categories by year. In the smallest size category the fraction of elections held where $p < 0.5$ is consistently less than 0.2. In contrast, in the largest size category more than half of elections held in each year since 1970 are predicted to have had values of $p < 0.5$.

In the context of the model of union organizing behavior, these results are consistent with the view that unions are willing to contest large elections where they have a lower probability of prevailing. But they also imply that the specific functional form of the beta distribution should not be taken too literally. As noted earlier, the probability of a prounion vote is *not* the probability of a union victory. Unions are relatively unlikely to win large elections where p is even a small amount below 0.5. The fact that the implied beta distribution for large elections has substantial mass at some distance below 0.5, where the union's probability of winning the election is very small, suggests, taking the model at face value, that unions derive very large value from organizing large units. Two alternatives are 1) that the beta distribution is not the correct distribution or 2) that the probability of a prounion vote can change after the union files for election (perhaps as a result of campaign activity). While beyond the scope of this study, the former possibilty can be investigated using the micro-data on election outcomes available after 1973. The latter possibility suggests a richer model of the union choice process but does not fundamentally alter the analysis.

Not withstanding these potential extensions, a beta-binomial model that constrains the changes over time in the mean of the distribution of p to be constant across size categories fits the win rates very well and can account for the increase in the size gap in win rates without invoking an explanation based on differential movements in the distribution of p across size categories.

7. Final Remarks

I established a set of facts regarding union election activity in the post-World War II period. These include 1) The quantity of election activity has fallen sharply and discontinuously since the mid-70's after increasing since between the mid-1950's and mid-1970's; 2) Union success in elections held has declined less sharply and essentially continuously over the entire period; 3) It has always been the case that unions have been less likely to win NLRB-supervised representation elections in large units than in small units; and 4) The size-gap in union success rates has increased substantially over the last forty years. I then developed a simple optimizing model of the union decision to hold a representation election that reconciles the first two facts and that directly implies the third.

I proposed two potential explanations for the fourth fact, that the size gap in union success rates has been growing over time. One is based on differential behavior by employers of different sizes as larger employers may be resisting union election activity more strongly than in the past. The second explanation is purely statistical and based on the idea that, as the probability that an individual votes in favor of union representation falls, the win rates will fall more in large elections than in small elections. In order to distinguish these explanations, I estimated three empirical models of election outcomes using aggregate data on NLRB elections outcomes over the 1952-98 time period in order to determine whether the simple statistical model can account for the size pattern of union win rates over time.

I find that a model constraining the mean difference in the pro-union vote probability between large and small units to be fixed over time but allowing for common time effects and random heterogeneity across elections in the pro-union vote probability accounts for the time-series patterns of win rates quite well.

The overall pattern of results implies that systematic union selection of targets for organization combined with the purely statistical explanation of the increase in the size gap in win rates can account for the four facts.

I have two concluding comments. First, differential behavior by employers in large and small establishments is not required to account for the facts, but neither is it ruled out. In particular, it is possible that large employers fight more strongly against union organization. In the context of the model, this would shift the distribution of θ (the probability of a union win) to the left in large units. As the selection model suggests, this would affect the quantity of election activity (by reducing it in large units), but the union win rate in large elections would not be affected. I do find that the quantity of elections in large units has declined, but this might simply reflect a diminution of the pool of viable large targets over time due to previous successful organization or to changes in the distribution of employment away from large establishments with workers interested in unions (perhaps a result of the long-term shift away from manufacturing employment).

Finally, the evidence is very clear that there is substantial heterogeneity across elections in the underlying probability of voting for union representation (p). This evidence comes from two sources. First, the binomial model simply implies too smooth a time path of pro-union vote shares over time (figure 12). This comes from the necessity to fit the win rates in large elections where small deviation of p from 0.5 imply dramatic movements in the win rate. Second, the estimates of the beta-binomial model imply that there is substantial heterogeneity (figure 15). Preliminary analysis of the distribution of votes shares from 1973 on using the micro-data (not presented here) also shows directly that there is substantial heterogeneity in p, and it implies that an even richer characterization of the distribution of p than a beta is likely necessary. Analysis of the structure of this distribution, how it has moved over time, and how it varies across size categories and other establishment characteristics will be the focus of future work.

8. References

Dickens, William T. "The Effect of Company Campaigns on Certification Elections: *Law and Reality* Once Again," *Industrial and Labor Relations Review* 36 (July 1983): pp. 560-575.

Dickens, William T. and Jonathan S. Leonard. "Accounting for the Decline in Union Membership," ' *Industrial and Labor Relations Review* 38 (April 1985): pp. 323-334.

Evans, Hastings, and Peacock. *Statistical Distributions*, New York, John Wiley & Sons, 1993.

Farber, Henry S. "The Extent of Unionization in the United States: Historical Trends and Prospects for the Future," in *Challenges and Choices Facing American Labor*, Thomas Kochan, ed. M.I.T. Press, 1985.

Farber, Henry S. "The Decline of Unionization in the United States: What Can be Learned from Recent Experience?," *Journal of Labor Economics* 8(1) January 1990: pp. S75-S105.

Farber, Henry S. and Alan B. Krueger. "Union Membership in the United States: The Decline Continues," in *Employee Representation: Alternatives and Future Directions*, Bruce Kaufman and Morris Kleiner, editors. Industrial Relations Research Association, 1993.

Freeman, Richard B. "Why are Unions Faring so Poorly in NLRB Representation Elections?," in *Challenges and Choices Facing American Labor*, Thomas Kochan, ed. M.I.T. Press, 1985.

Freeman, Richard B. and James L. Medoff. *What do Unions Do?*, New York, Basic Books, 1984.

Gould, William B. IV. *Agenda for Reform: The Future of Employment Relationships and the Law*, Cambridge MA, MIT Press, 1993.

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

Lalonde, Robert J. and Bernard D. Meltzer. "Hard Times for Unions: Another Look at the Significance of Employer Illegalities," *The University of Chicago Law Review* 58 (1991): pp. 953-1014.

Levy, Paul A. "The Unidimensional Perspective of the Reagan Labor Board," Rutgers Law Journal 269 (1985).

Riddell, W. Craig. "Unionization in Canada and the United States: A Tale of Two Countries," in *Small Differences that Matter: Labor Markets and Income Maintainance in Canada and the United States*, David Card and Richard B. Freeman, editors. Chicago, University of Chicago Press, 1993.

Weiler, Paul C. "Promises to Keep: Securing Workers' Rights to Self-organization under the NLRA," *Harvard Law Review* 96 (1983): pp. 1769-1827.

Weiler, Paul C. Governing the Workplace: The Future of Labor and Employment Law, Cambridge MA, Harvard University Press, 1990.

Weiler, Paul C. "Hard Times for Unions: Challenging Times for Scholars." *The University of Chicago Law Review* 58 (1991): pp. 1015-1032.

Western, Bruce. *Between Class and Market: Postwar Unionization in the Capitalist Democracies*, Princeton NJ, Princeton University Press, 1997.

</ref_section>

TABLE 1: Union Win Rates NLRB Representation Elections
by Size Category and Year

Year	1-9 Votes	10-19 Votes	20-49 Votes	50-99 Votes	≥100 Votes
1952	0.78	0.74	0.72	0.71	0.69
1953	0.77	0.72	0.70	0.67	0.72
1954	0.71	0.66	0.64	0.64	0.61
1955	0.75	0.70	0.66	0.62	0.66
1956	0.70	0.68	0.67	0.61	0.60
1957	0.71	0.68	0.63	0.57	0.51
1958	0.71	0.64	0.59	0.56	0.55
1959	0.68	0.64	0.63	0.61	0.56
1960	0.66	0.62	0.55	0.52	0.56
1961	0.65	0.60	0.53	0.52	0.46
1962	0.65	0.59	0.58	0.55	0.51
1963	0.65	0.63	0.58	0.54	0.51
1964	0.68	0.60	0.56	0.52	0.47
1965	0.67	0.63	0.60	0.54	0.55
1966	0.69	0.64	0.60	0.56	0.54
1967	0.67	0.63	0.59	0.55	0.50
1968	0.63	0.63	0.58	0.52	0.49
1969	0.63	0.58	0.55	0.52	0.46
1970	0.66	0.58	0.54	0.53	0.46
1971	0.63	0.59	0.53	0.47	0.41
1972	0.63	0.60	0.55	0.45	0.42
1973	0.60	0.55	0.50	0.48	0.41
1974	0.62	0.55	0.48	0.45	0.36
1975	0.58	0.53	0.48	0.42	0.36
1976	0.59	0.56	0.47	0.44	0.34
1977	0.58	0.52	0.47	0.44	0.36
1978	0.59	0.50	0.47	0.42	0.31
1979	0.59	0.52	0.48	0.41	0.31
1980	0.57	0.50	0.48	0.44	0.32
1981	0.55	0.51	0.46	0.40	0.31
1982	0.52	0.50	0.43	0.36	0.31
1983	0.54	0.50	0.43	0.42	0.38
1984	0.54	0.52	0.45	0.41	0.39
1985	0.57	0.50	0.44	0.42	0.37
1986	0.58	0.52	0.44	0.43	0.32
1987	0.58	0.54	0.46	0.43	0.35
1988	0.58	0.54	0.48	0.46	0.38
1989	0.57	0.55	0.48	0.46	0.36
1990	0.61	0.54	0.51	0.44	0.36
1991	0.57	0.52	0.46	0.40	0.32
1992	0.61	0.55	0.47	0.43	0.34
1993	0.57	0.55	0.51	0.44	0.39
1994	0.61	0.56	0.46	0.45	0.39
1995	0.61	0.55	0.48	0.46	0.36
1996	0.65	0.51	0.47	0.44	0.34
1997	0.62	0.51	0.46	0.46	0.37
1998	0.62	0.57	0.52	0.48	0.37
Average	0.63	0.58	0.53	0.49	0.44

TABLE 2: Average Pro-Union Vote Share in NLRB Representation Elections by Size Category and Year

Year	1-9 Votes	10-19 Votes	20-49 Votes	50-99 Votes	≥100 Votes
1973	0.62	0.57	0.53	0.53	0.50
1975	0.60	0.55	0.52	0.49	0.48
1976	0.61	0.55	0.51	0.51	0.47
1977	0.60	0.54	0.51	0.50	0.47
1978	0.60	0.52	0.51	0.49	0.45
1979	0.61	0.55	0.51	0.48	0.46
1980	0.60	0.52	0.51	0.50	0.46
1981	0.58	0.53	0.50	0.48	0.46
1982	0.55	0.52	0.50	0.46	0.46
1983	0.57	0.53	0.49	0.48	0.50
1984	0.57	0.53	0.50	0.48	0.49
1985	0.60	0.53	0.49	0.49	0.49
1986	0.61	0.54	0.49	0.49	0.46
1987	0.60	0.55	0.51	0.49	0.48
1988	0.61	0.56	0.53	0.51	0.48
1989	0.61	0.56	0.53	0.51	0.47
1990	0.62	0.56	0.54	0.51	0.48
1991	0.60	0.53	0.51	0.48	0.46
1992	0.64	0.57	0.52	0.50	0.47
1993	0.60	0.57	0.53	0.50	0.49
1994	0.64	0.57	0.51	0.51	0.48
1995	0.64	0.56	0.51	0.50	0.47
1996	0.67	0.54	0.51	0.49	0.46
1997	0.63	0.55	0.51	0.51	0.47
1998	0.63	0.58	0.54	0.51	0.48
Average	0.60	0.55	0.51	0.50	0.47

TABLE 3: Estimates of Constrained Models of Union Vote Share Based on Aggregate Win Rate Data

Variable	Binomial	Beta Mixture
Constant	0.660	0.717
	(0.0004)	(0.0030)
1-9 voters	---	---
10-19 Voters	-0.078	-0.077
	(0.0005)	(0.0015)
20-49 Voters	-0.111	-0.113
	(0.0005)	(0.0013)
50-99 Voters	-0.128	-0.133
	(0.0006)	(0.0014)
$\geq$ 100 Voters	-0.133	-0.153
	(0.0004)	(0.0017)
Year Effects	Yes	Yes
α	---	15.72
		(1.36)
Log-L	-1237.3	-926.8

These are maximum-likelihood estimates of the probability of the union win rate using aggregate data on annual win rates in five size categories from 1952 through 1998. The model in the first column is a binomial model assuming a fixed probability of a pro-union vote within size categories at a point in time. The model in the second column is a binomoial model which assumes that the vote share within size categories at a point in time have a beta distribution. The parameter α controls the variance of this beta distribution as described in the text. All models also include a set of 46 dummy variables for fiscal year. The numbers in parentheses are standard errors.

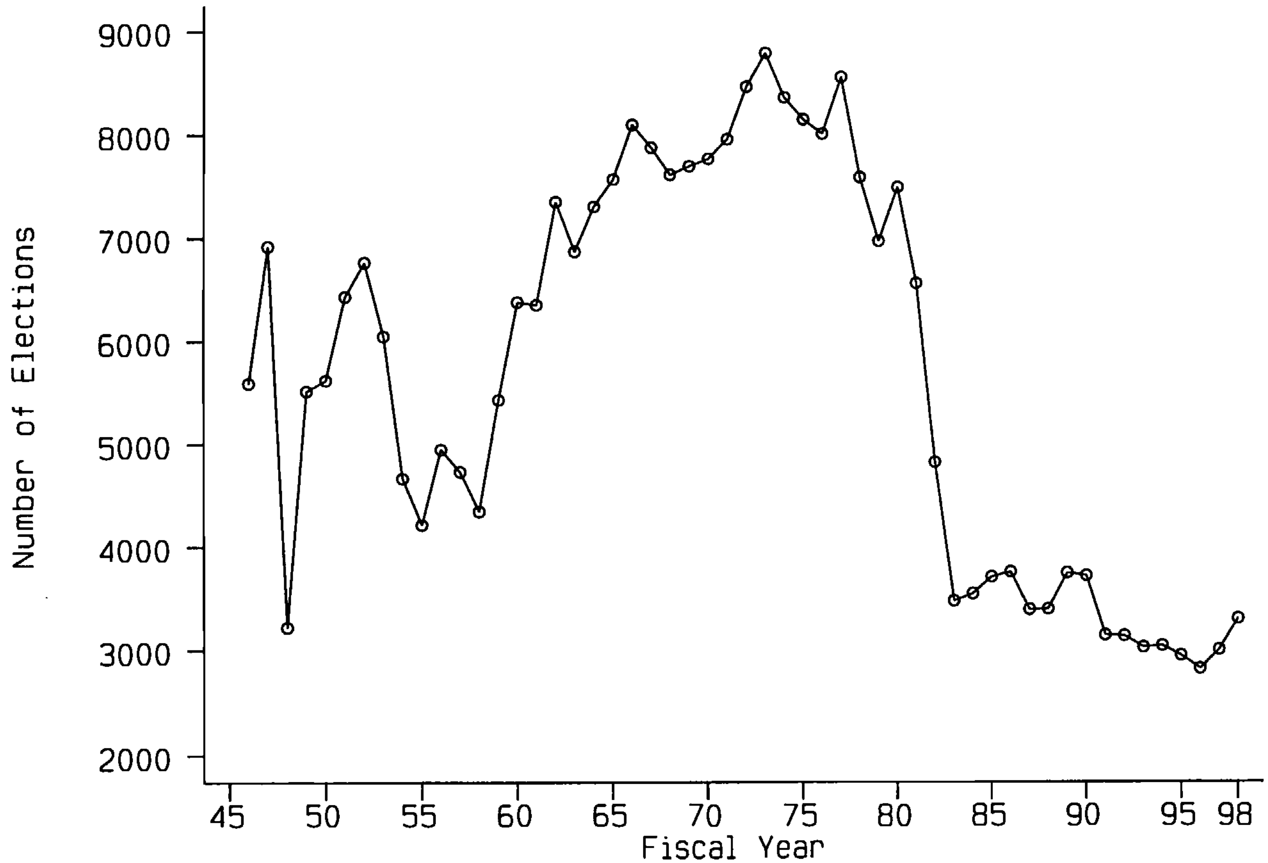


Figure 1 - Number of NLRB Certification Elections

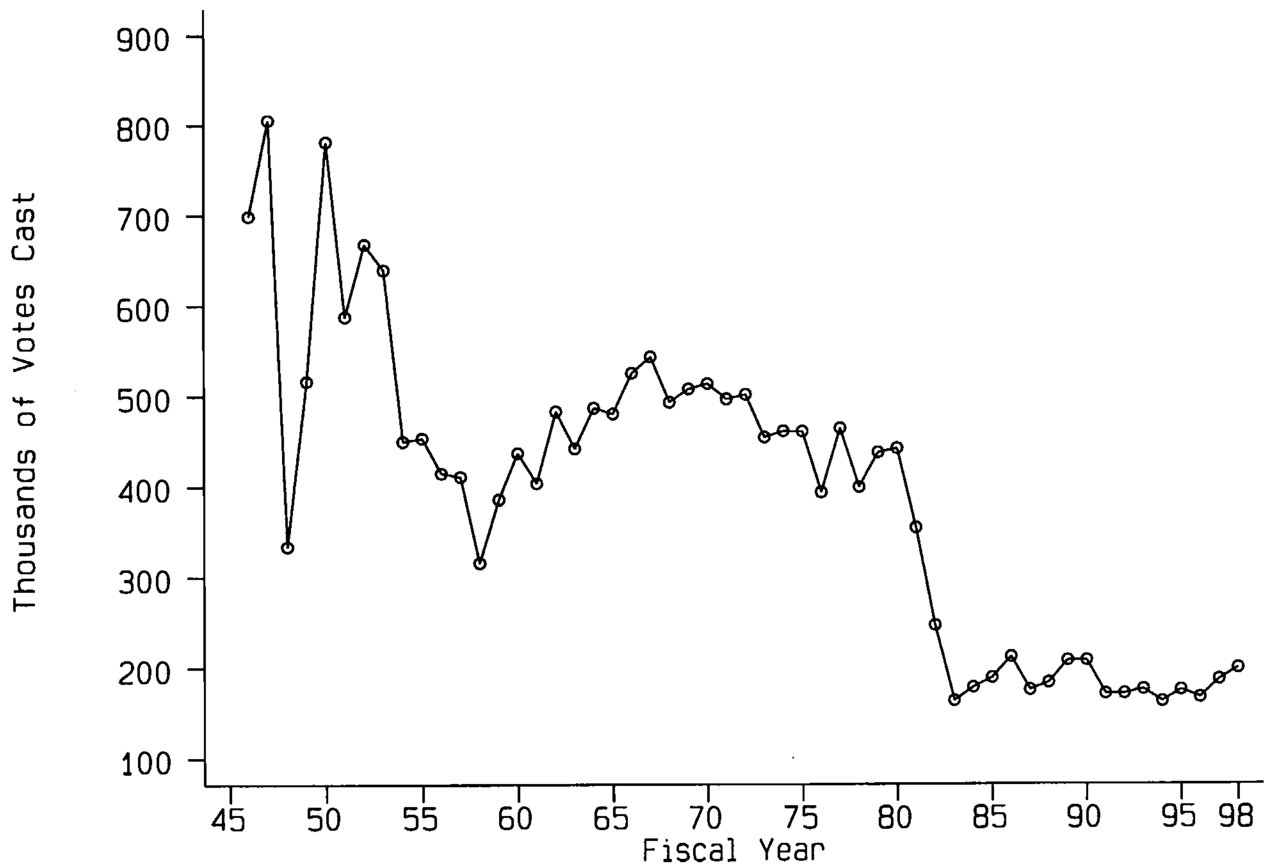


Figure 2 - Total Votes Cast in Elections (thousands)

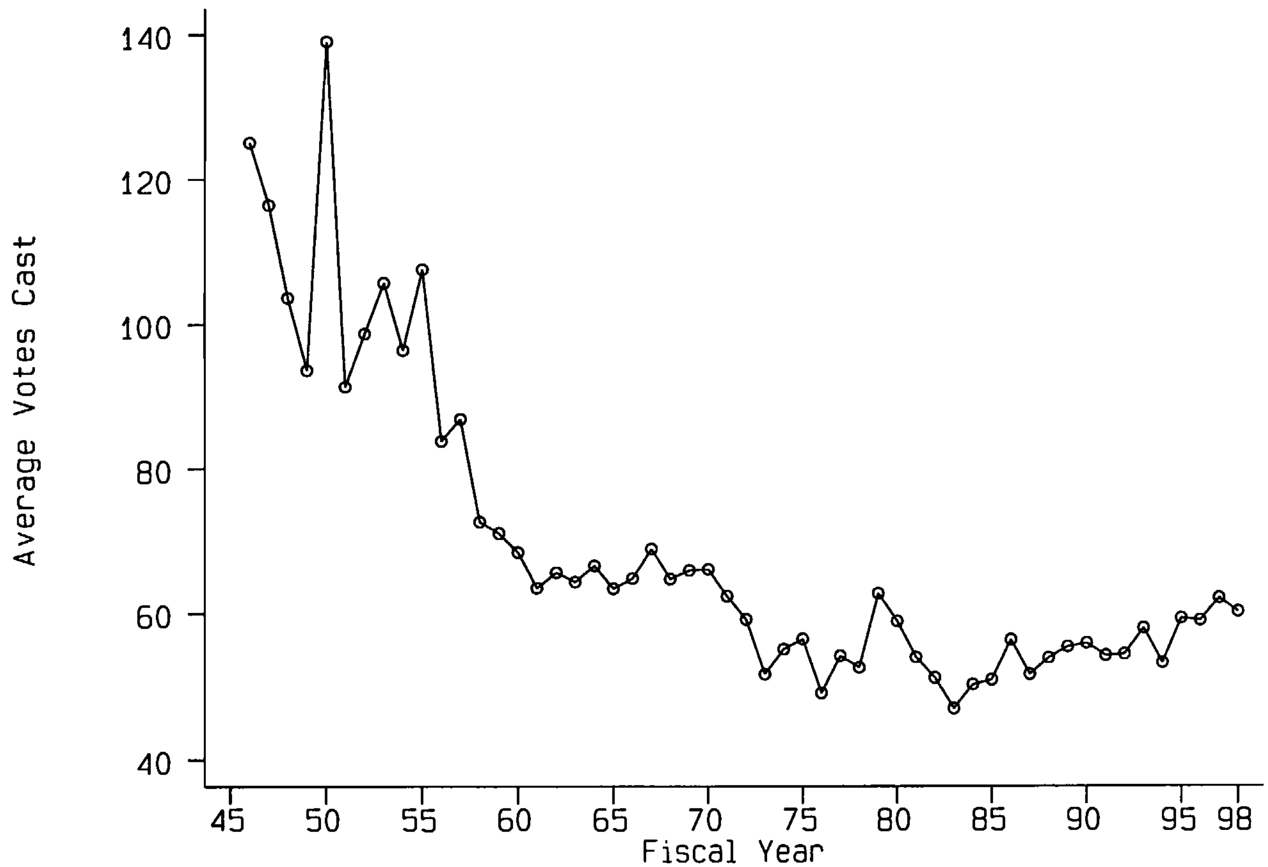


Figure 3 - Average Votes Cast per Election

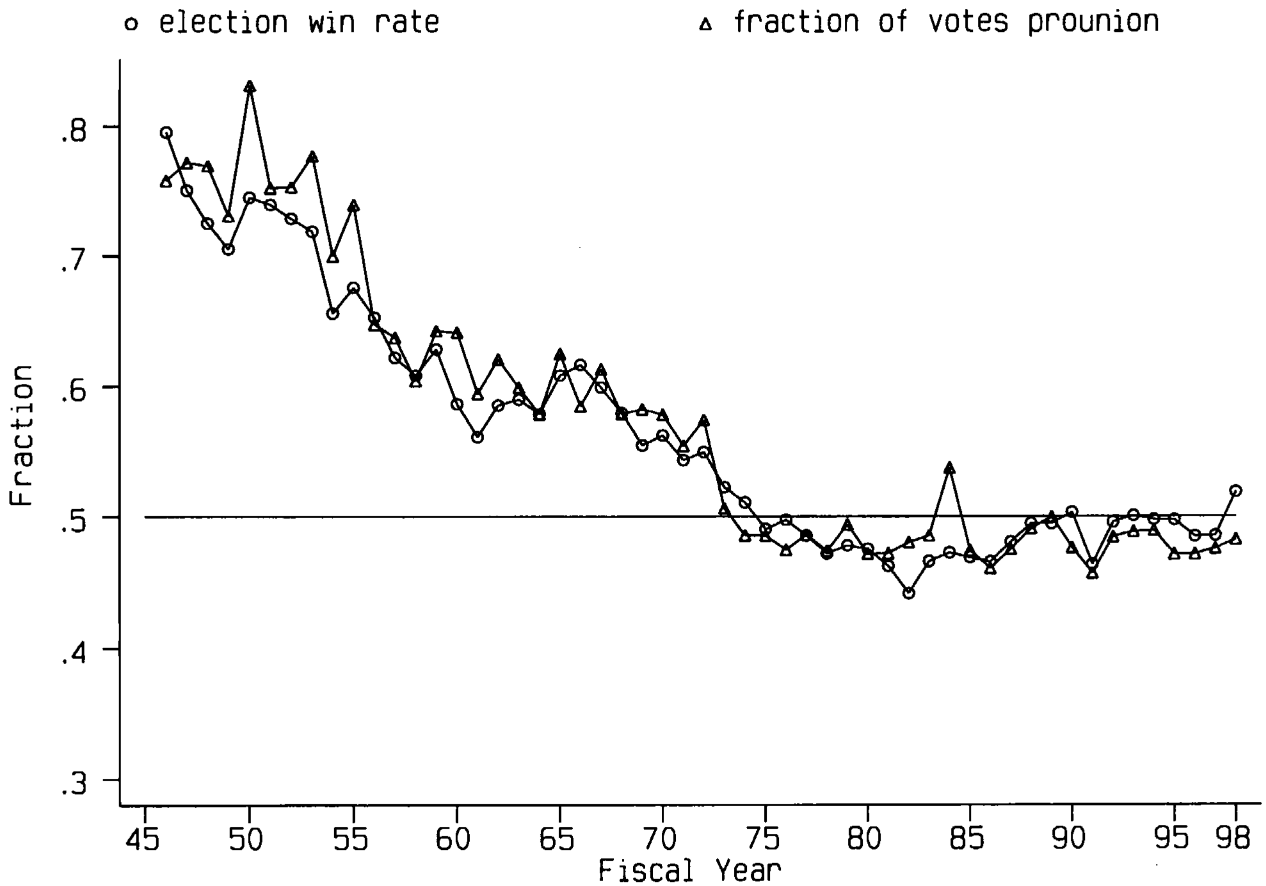


Figure 4 - Union Win Rate and Vote Share

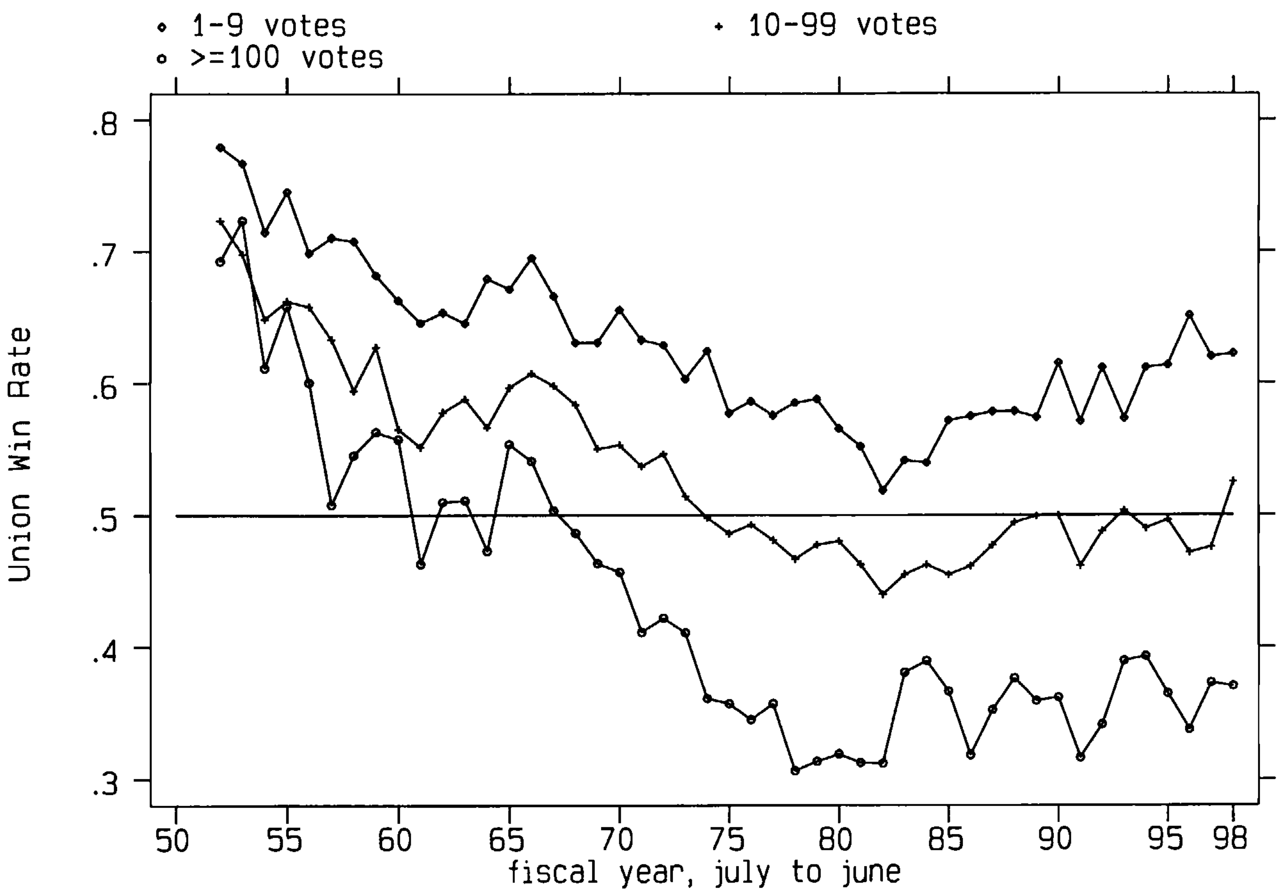


Figure 5 - Fraction Elections Won by Union

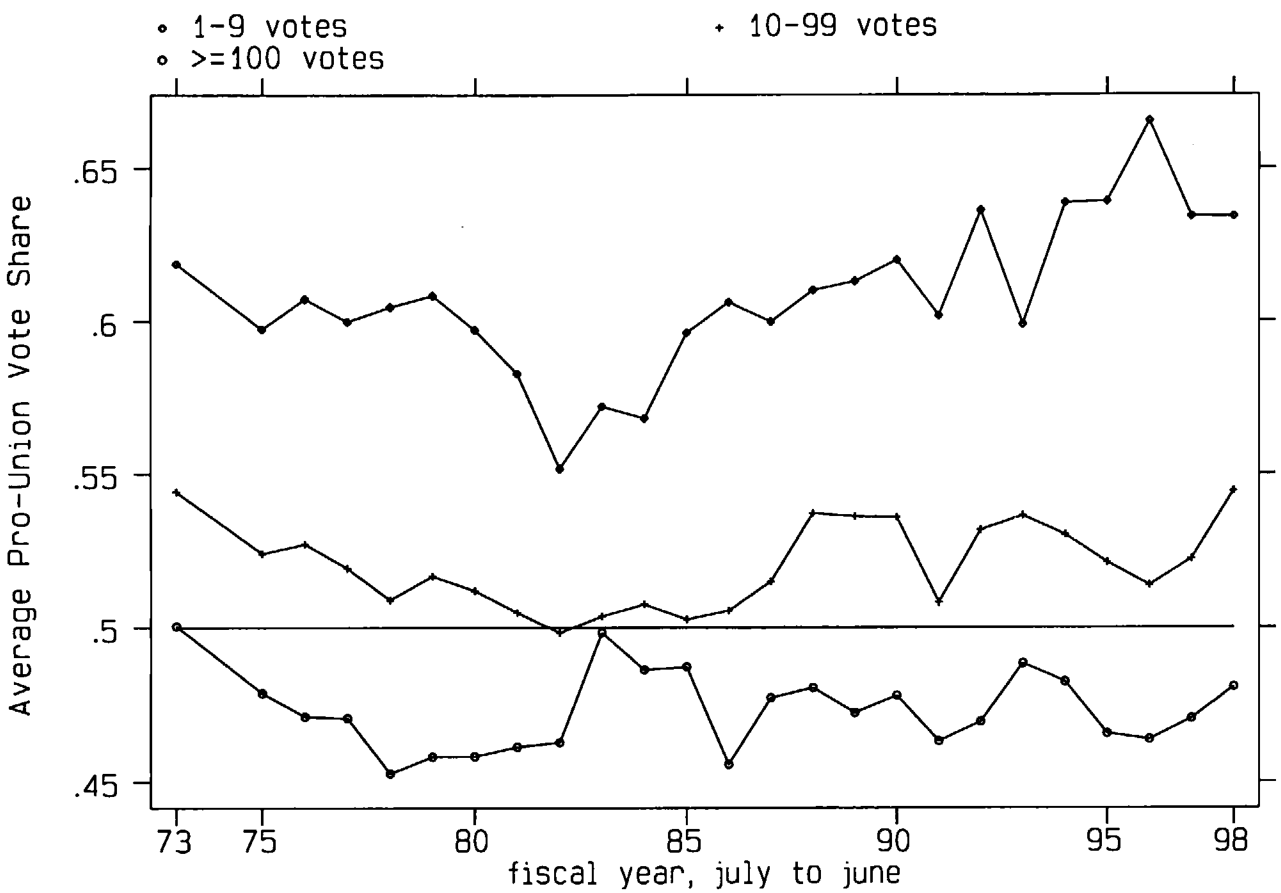


Figure 6 - Pro-Union Vote Share, Election Average

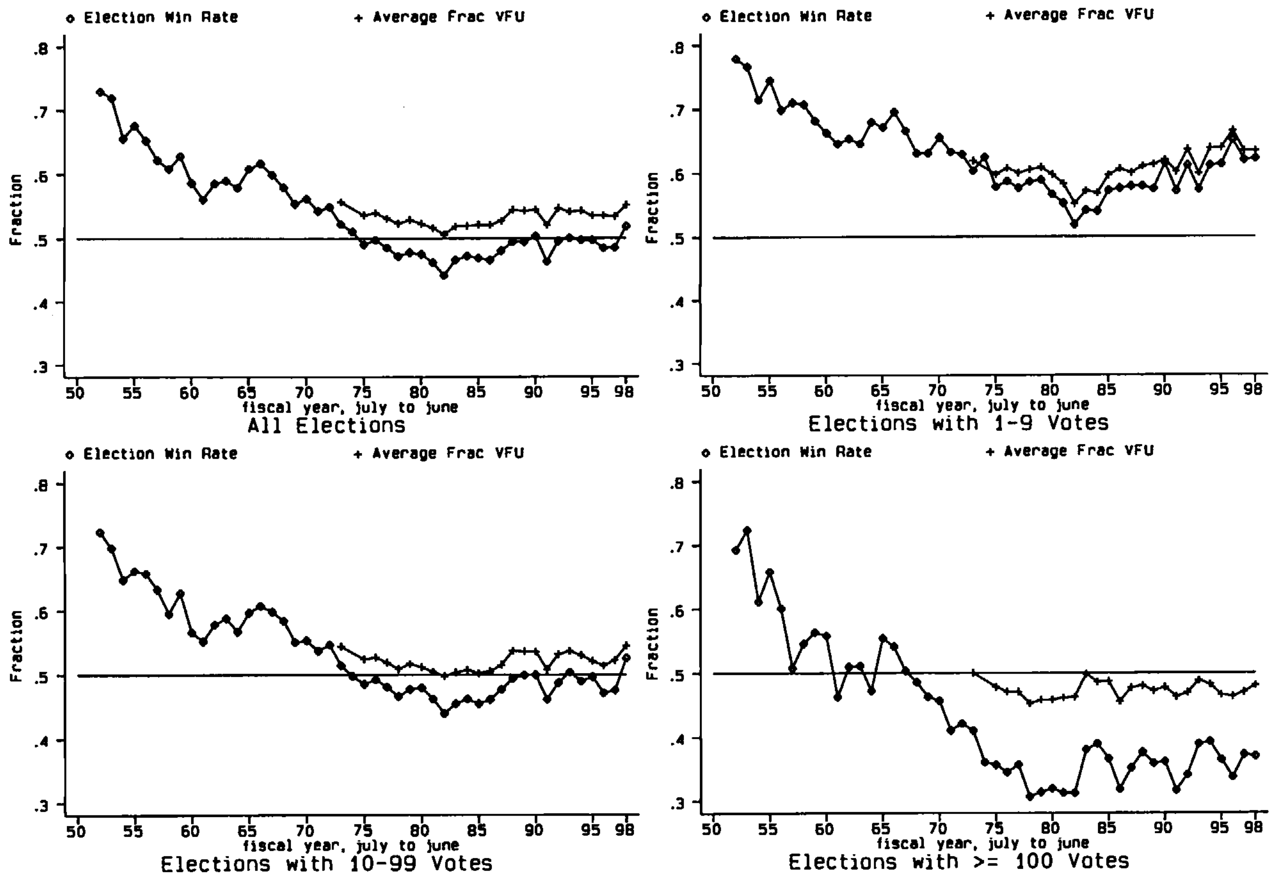
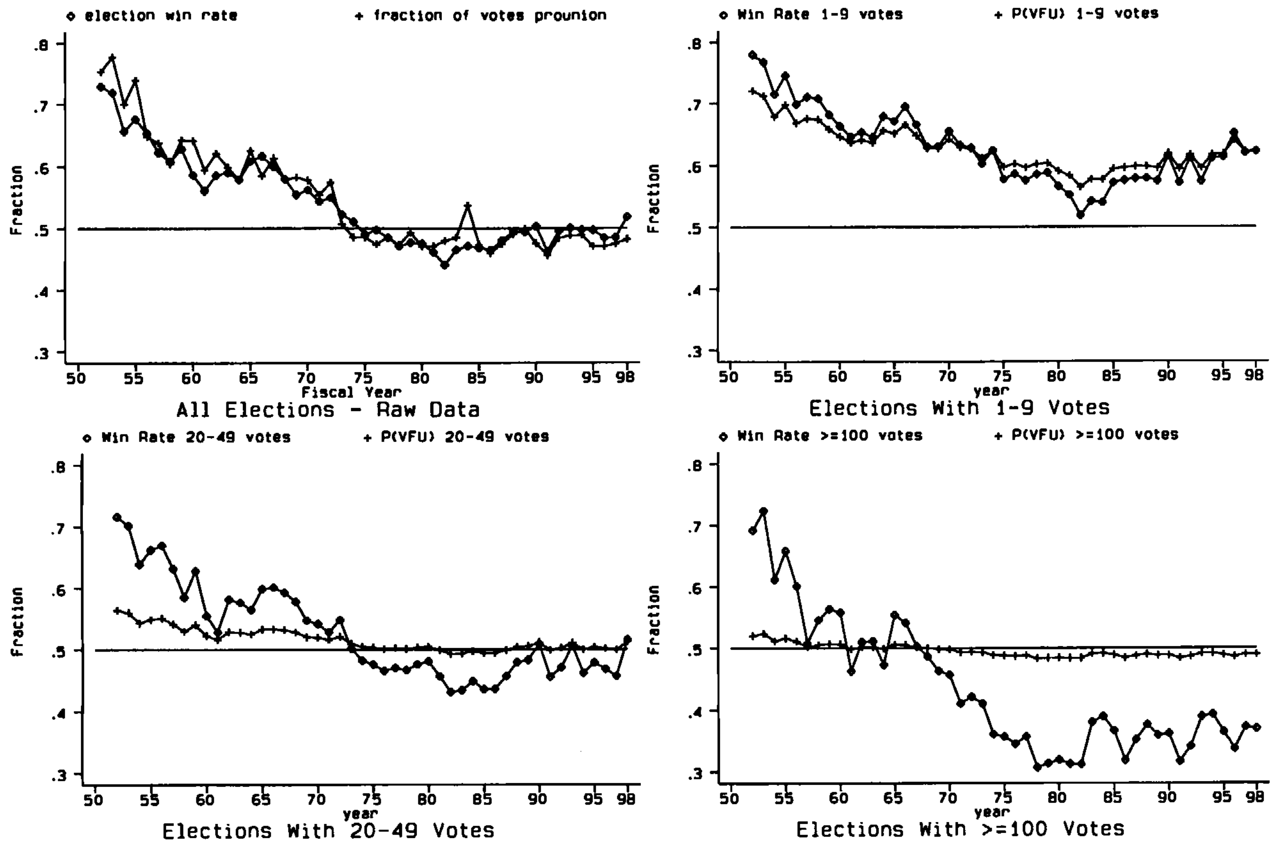


Figure 7 - Union Win Rate and Average Vote Share



Saturated Binomial

Figure 8 - Union Win Rate and Predicted Vote Share

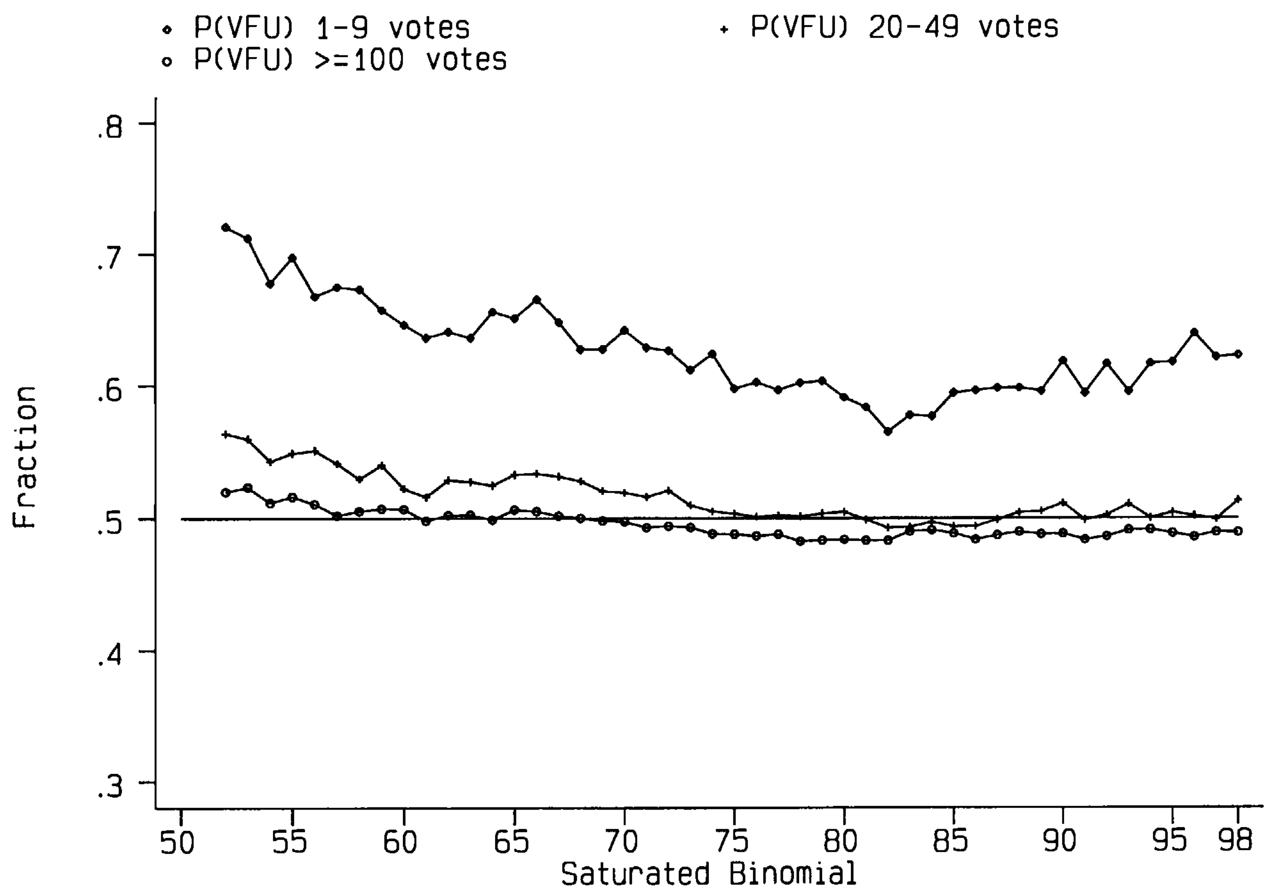


Figure 9 - Predicted Vote Share

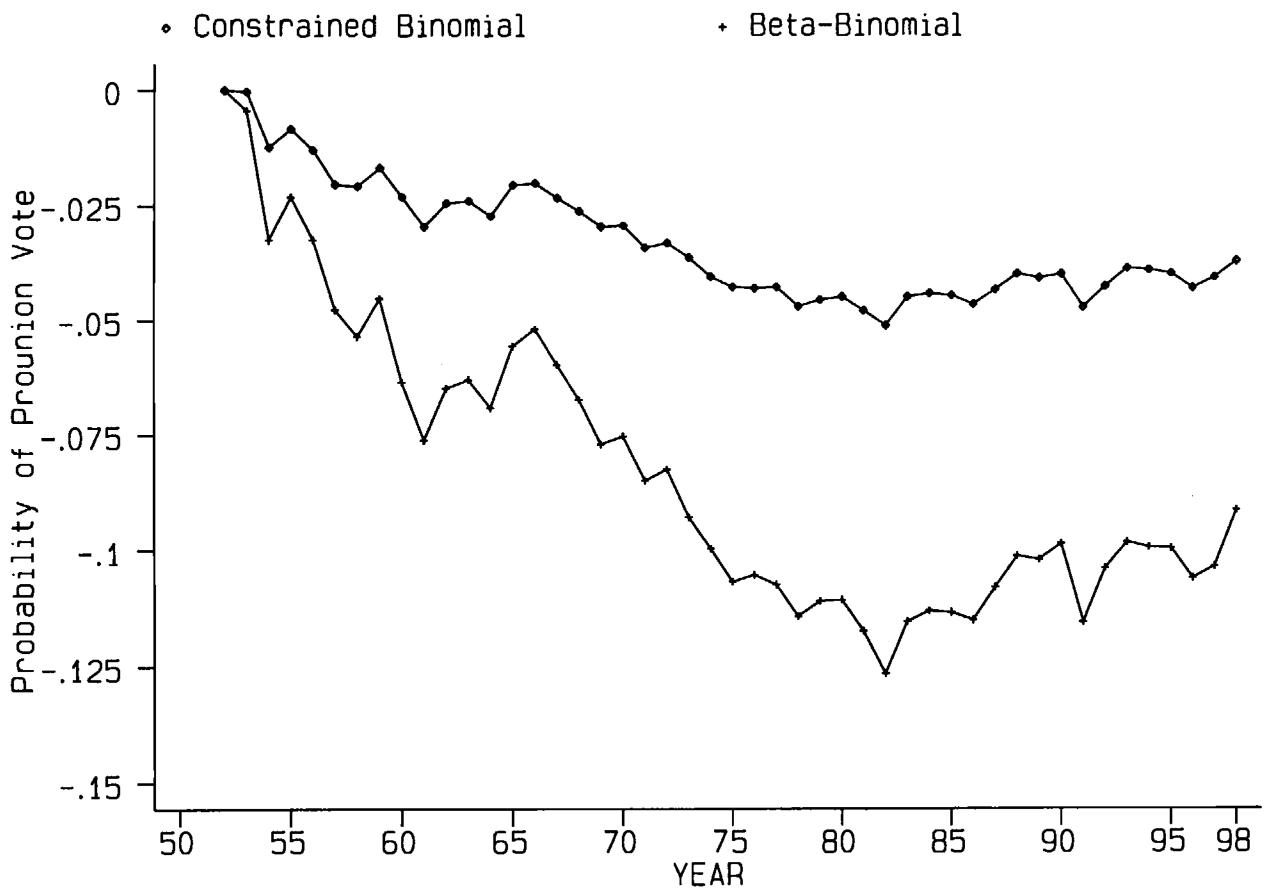
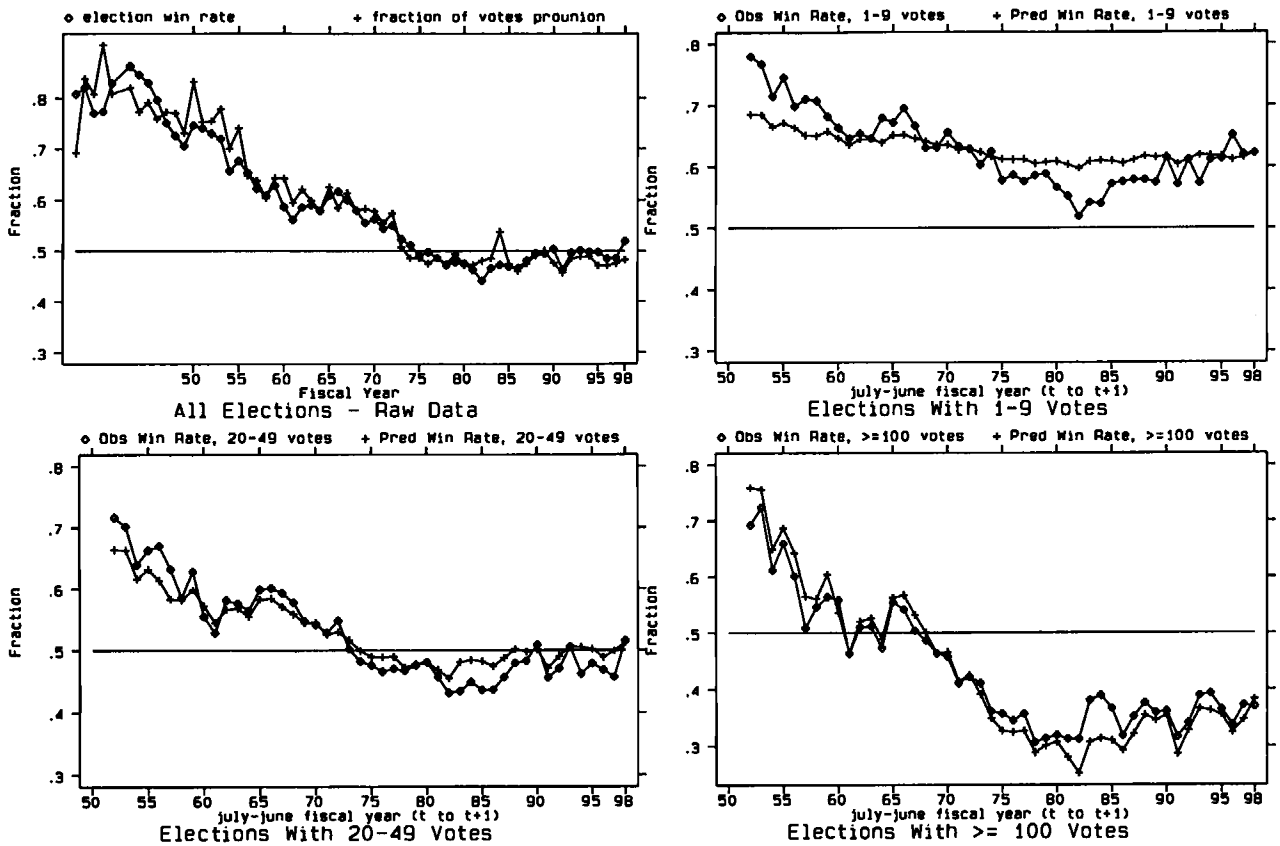
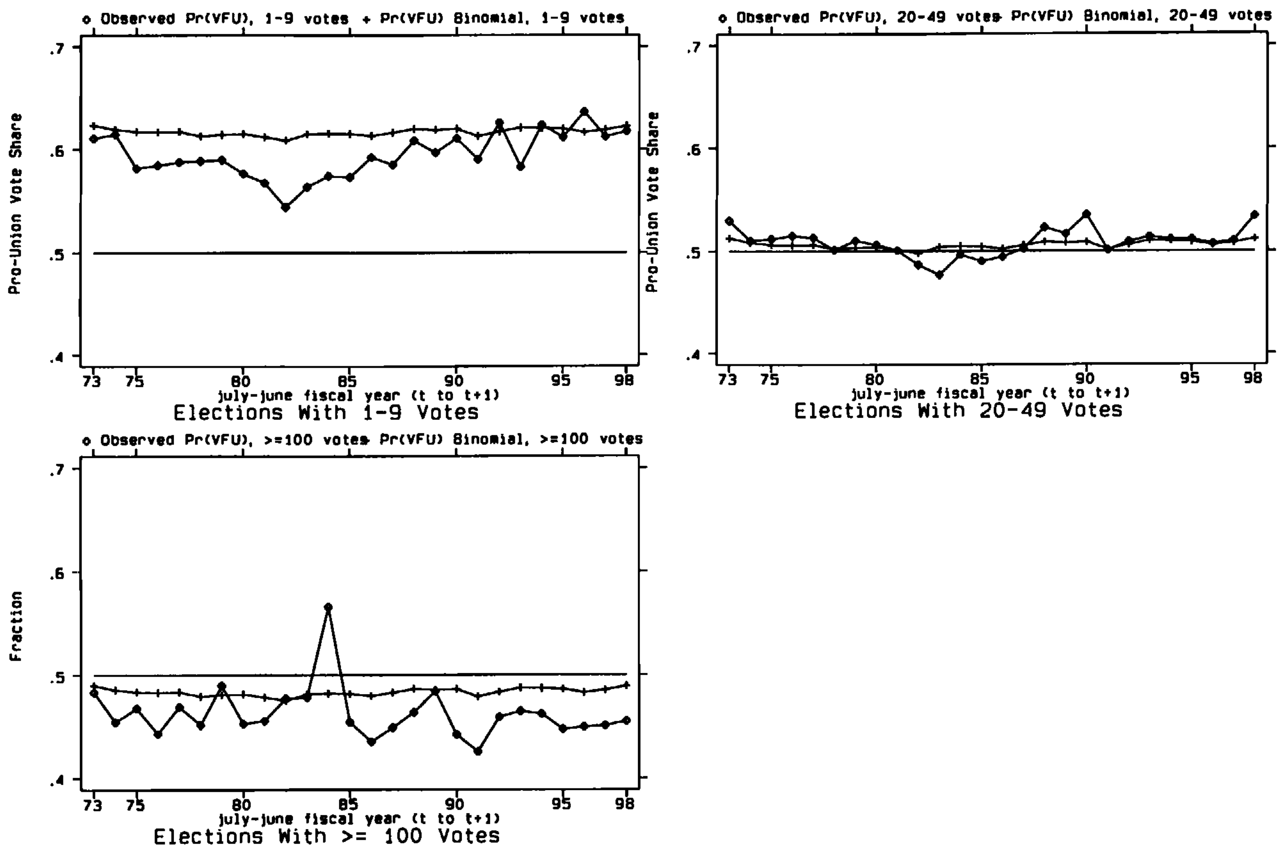


Figure 10 - Year Effects in Prounion Vote Share



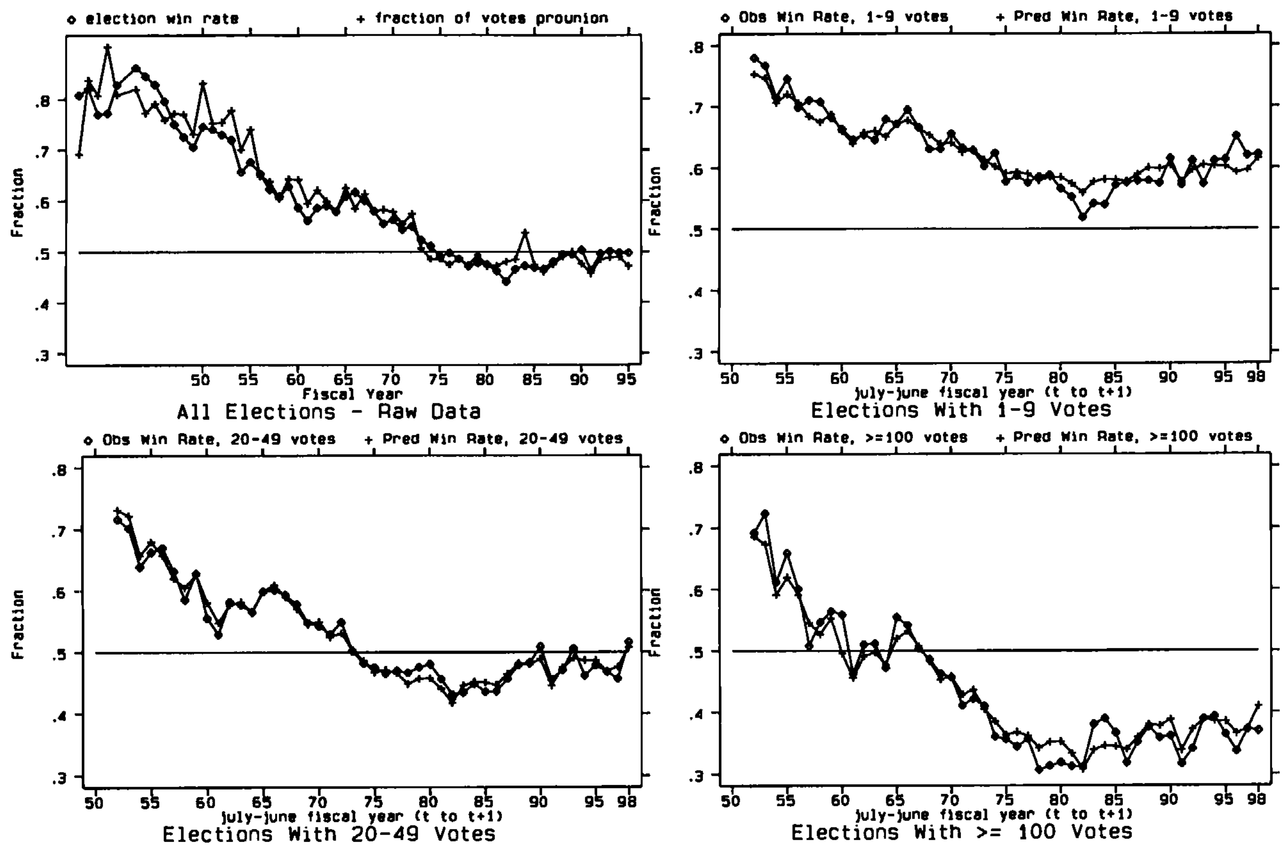
Constrained Binomial

Figure 11 - Actual and Predicted Union Win Rate



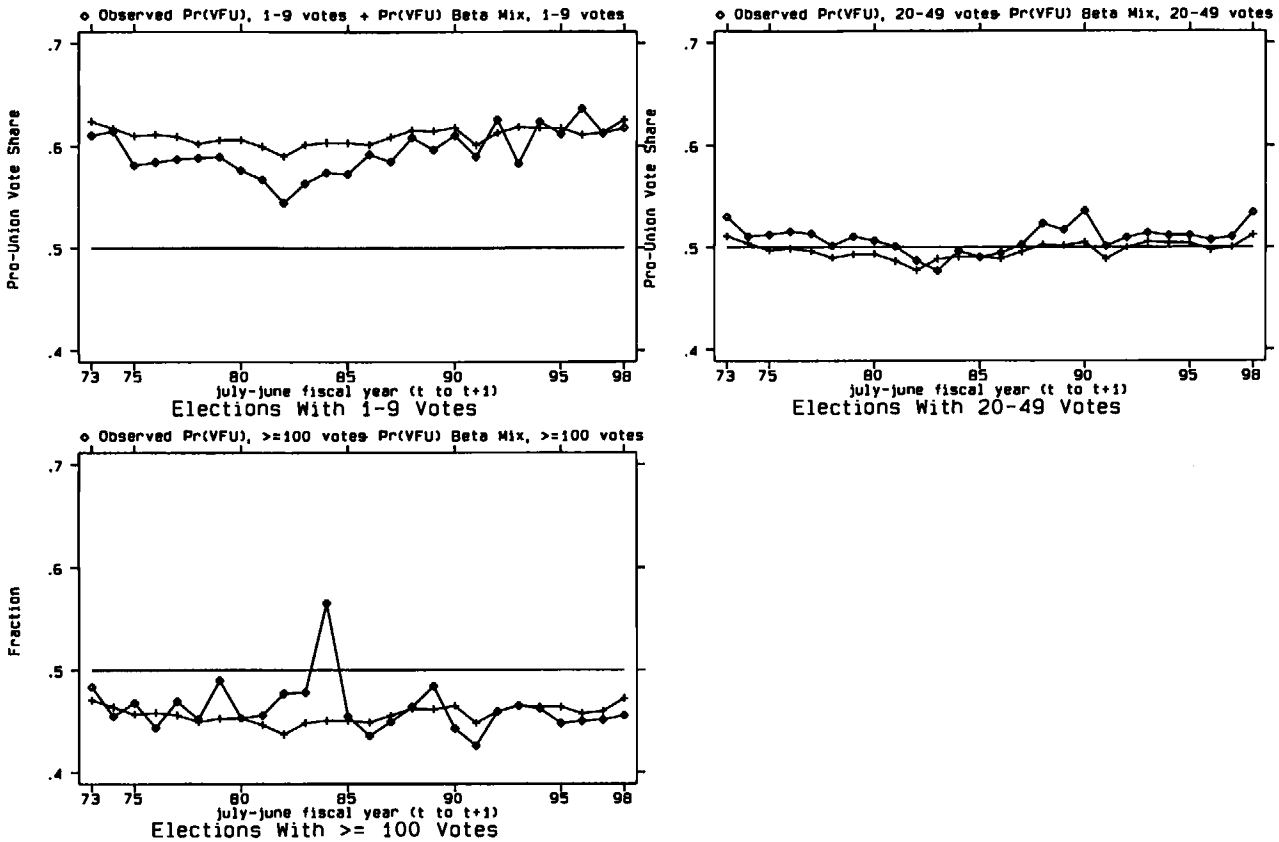
Constrained Binomial

Figure 12 - Actual and Predicted Pro-Union Vote Shares



Beta-Binomial

Figure 13 - Actual and Predicted Union Win Rate



Beta Binomial

Figure 14 - Actual and Predicted Pro-Union Vote Shares

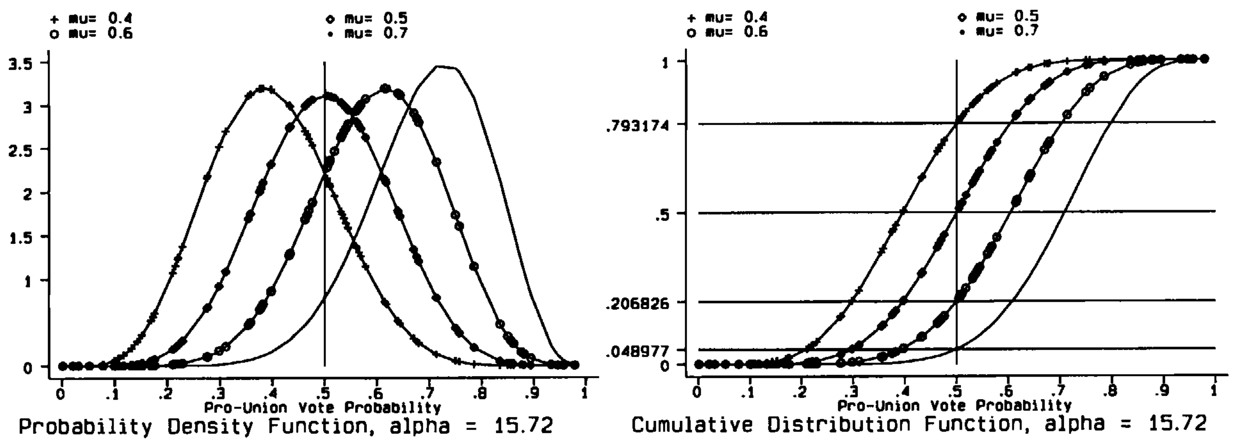


Figure 15 - Beta Distribution, Pro-Union Vote Probability

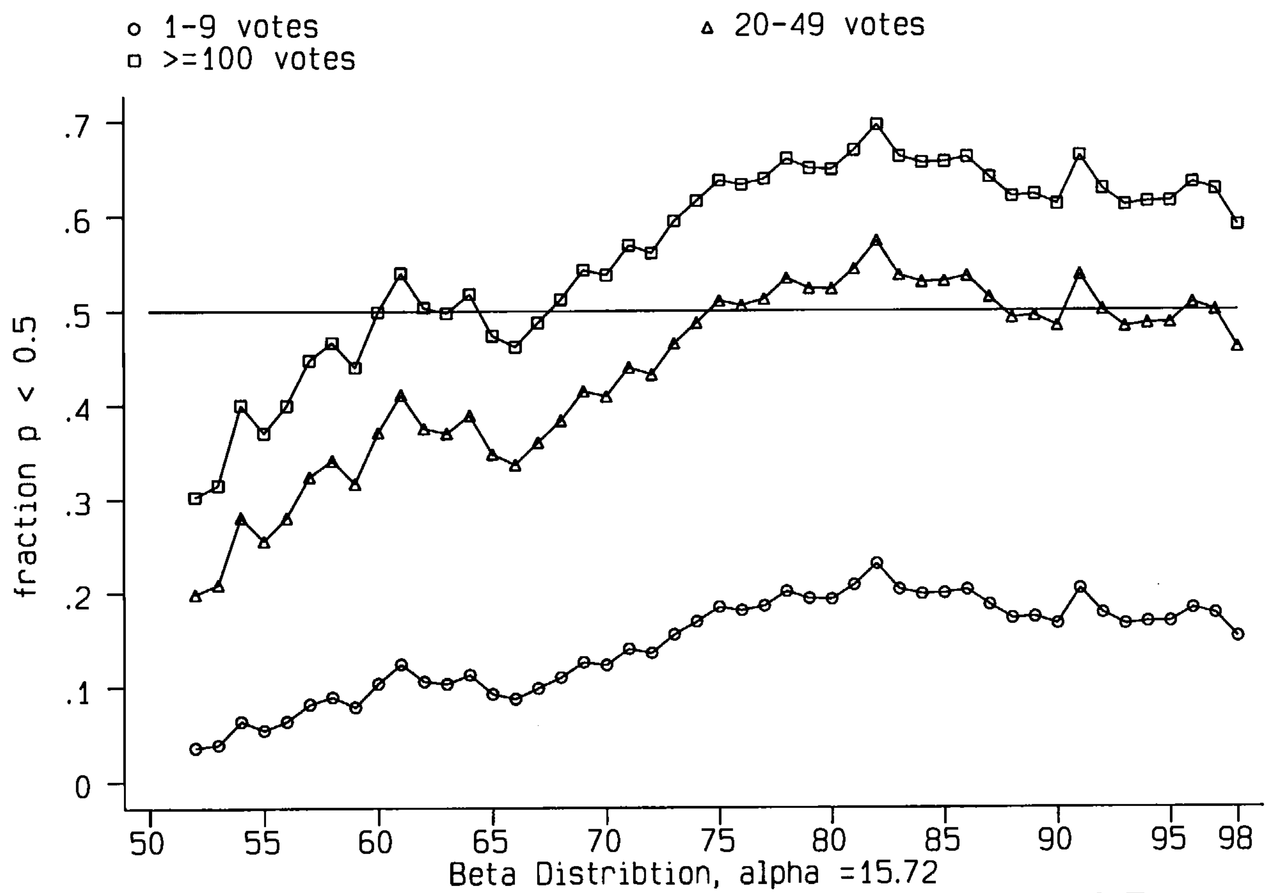


Figure 16 - Pred Frac Elections with p < 0.5