UNIVERSITY OF CALIFORNIA,
IRVINE

# Essays on Collective Bargaining and Strikes

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Kyung nok Chun

Dissertation Committee:
Professor Stergios Skaperdas, Chair
Professor Michelle Garfinkel
Professor Matthew Freedman

2021

# DEDICATION

To Yesong and Siwon, the light of my life

# TABLE OF CONTENTS

# LIST OF FIGURES

Page

# LIST OF TABLES

Page

# ACKNOWLEDGMENTS

I thank first and foremost my partner Yesong, who has had to stay home to raise a young child with special needs in a country that offers neither affordable child care nor the chance for 'dependents' of international students to work for pay. Not a single paragraph of this dissertation would have been possible without her superhuman sacrifice and support. I hope she never gets to read it and find out what it was all for.

I thank my extended family for their moral and financial support, which kept me going through rough patches.

I thank God for letting my advisor Stergios Skaperdas survive his heart attack during my second year in the PhD program. I highly doubt I would have made it this far without his infinitely patient, compassionate mentorship. Stergios has inspired me through his own example and encouragement/subtle manipulation to tackle unconventional problems, and my intellectual journey has been all the richer for it. Also Thanksgiving dinner at his place stands out as one of the warmest memories Yesong and I have of Irvine.

I thank my committee members Michelle Garfinkel and Matthew Freedman for somehow always finding time to read through my confused manuscripts and provide thoughtful feedback. Michelle has been instrumental in my conversion from the empirical to the theoretical side of economics. Matthew's comments on some of my earlier work have directly inspired the idea for the third chapter of this dissertation.

I am also grateful to faculty members outside of my committee who at various points helped me develop as a researcher. I would like to acknowledge in particular David Neumark, who pushed me to think hard about the theory motivating the empirics; and Emily Owens, who impressed upon me the value of socially conscious scholarship.

Special thanks are also due to my upper-year colleagues Zachary Schaller, Byunggeor Moon, and Nathan Cisneros, who served in my informal committee of peer advisors. Zach's research on labor unions has been an example for me to emulate, and his numerous feedback and help with my own work has been of tremendous value. Byunggeor helped me navigate the numerous challenges I was to encounter throughout my stay in the program, and went over and beyond to prepare me for interviews in a terrible job market year – he is solely responsible for my failure to secure a job. Nathan is also a researcher of labor unions with whom I've had lots of thought-provoking conversations, but more than that has been a friend who genuinely empathized with what my family and I was going through.

Two other colleagues of mine deserve special mention for having had the misfortune of being my closest friends and as a result went out of their way to assist my family's rushed exodus during the pandemic: Aria Golestani and Yoonjung Kim.

Last but not least, I thank Wonil Kim for being a genuine friend and a comforting presence through some of my darkest moments.

# VITA

## Kyung nok Chun

### EDUCATION

**Doctor of Philosophy in Economics** **2021**
University of California, Irvine *Irvine, CA, USA*

**Master of Arts in Economics** **2016**
University of Texas at Austin *Austin, TX, USA*

**Master of Arts in Translation and Interpretation** **2008**
Hankuk University of Foreign Studies *Seoul, South Korea*

**Bachelor of Law** **2006**
Korea University *Seoul, South Korea*

### REFEREED JOURNAL PUBLICATIONS

**Why are there strikes? (With Stergios Skaperdas and Zachary Schaller)** **2020**
*Revue d'économie politique*, 130 (6), 929-956

### TEACHING EXPERIENCE

**Teaching Assistant** **2016–2021**
University of California, Irvine *Irvine, California*

### WORK EXPERIENCE

**English Editor** **Oct 2012 – Jun 2015**
The Export-Import Bank of Korea

**In-house Translator** **Dec 2008 – May 2012**
The Boston Consulting Group Seoul Office

**International Business News Correspondent** **Dec 2007 – Aug 2008**
OBS

# ABSTRACT OF THE DISSERTATION

Essays on Collective Bargaining and Strikes

By

Kyung nok Chun

Doctor of Philosophy in Economics

University of California, Irvine, 2021

Professor Stergios Skaperdas, Chair

Labor unions are a controversial and relatively little understood species of organization. While empirical research on the effects of unions on various labor market outcomes proliferated in the 1970s and 80s, theoretical understanding of the behavior of unions remains at its early stages. My dissertation aims to advance such understanding.

The first chapter delves into the rationale behind strikes, the ultimate source of a union's bargaining power and therefore the basis of its entire existence. The very occurrence of strikes has been challenging for standard economic theory to explain: given their manifest inefficiency, a mutually beneficial bargain should always be available such that strikes only serve as a weapon of non-use. Existing models of strikes tend to overcome this problem by assuming some form of private information. Borrowing insights developed in the field of conflict economics, I show that strikes can occur even in the absence of uncertainty (about the adversary's strength, costs, or reservation payoffs) once we consider the effect of the outcome of a strike on the relative strategic positions of the union and the firm in future interactions. I also show that shifts in the institutional environment in favor of one side can trigger strikes motivated by a desire to disrupt the existing balance of power, and use this framework to analyze strike waves observed in the United States since the late 19th Century.

The second chapter focuses on a special aspect of the internal politics of unions, namely the

tension between the union leadership and the rank and file that has been regularly noted by labor movement scholars and practitioners. I develop a theoretical model demonstrating that when union leaders maximize the size of their organization rather than the welfare of the workers they represent, they can seek a settlement with the employer that is poorer than what the workers can expect to obtain through a strike. This result agrees well with the historical pattern of union officials often acting as a conservative break against more militant rank-and-file demands, and formalizes some insights present in Michels (1915)'s famed Iron Law of Oligarchy. In an extension of the baseline model, I show that union security agreements can amplify incentives for the union leadership to under-represent the interests of the workers.

Right-to-work (RTW) laws prohibit union security agreements in many U.S. states, and are frequently portrayed as a mortal threat to unionism in the country. It is claimed that by encouraging free-riding on union services, these laws undermine the bargaining power and viability of unions. The third chapter tests the substance of such claims by evaluating the impacts of RTW laws passed in six U.S. states in the 21$^{st}$ Century. Using a mix of empirical methods including difference-in-differences, event studies, and synthetic controls, I find evidence that in the private sector, RTW laws decrease union coverage by more than 10 percent, all else equal. Surprisingly, I find RTW laws to have only a small and insignificant effect on free-riding behavior as measured by the share of unionized workers who are nonmembers. Moreover, union formation through NLRB-administered elections do not appear to be adversely affected by RTW. I also find evidence that RTW legislation increases union wage differentials by up to five percentage points, which is suggestive of a change in union bargaining behavior. Separately evaluating the effects of the 2018 U.S. Supreme Court decision in *Janus v. AFSCME*, which effectively made the entire U.S. public sector right-to-work, I find union coverage in the affected states to have changed little, and union wage differentials to have increased by more than five percentage points. Some of these findings are quite novel, and challenge conventional assumptions about how RTW laws impact unions.

# Chapter 1

# Strikes Under Complete Information: How Institutional Bias and Enduring Returns to Victory can Spark Labor Conflict

## 1.1 Introduction

After nearly three decades of near-continual decline, work stoppages in the U.S. are once again on the rise ever since the stunning victory of the West Virginia teachers' strike in early 2018. Across the world massive protest movements from Sudan to Hong Kong, Chile, and France continue to erupt, often involving trade unions (or professional associations) calling a strike.

Despite the ongoing significance of strikes, economists' understanding of the phenomenon has made little headway since the heyday of research on trade unions in the 1980s. Nearly all existing models of union-firm bargaining explain the occurrence of strikes by assuming some form of information asymmetry, the actual import of which is difficult to establish. No consensus has yet emerged on why strikes occur despite their inefficiency.

This chapter draws on recent theoretical advances in political economy (and conflict economics in particular) to develop a model of strikes under complete information. It will be shown that firms or unions can optimally choose to strike even when the size of the surplus to be bargained over, the objectives of the negotiating parties, and their win probabilities in the event of a strike are perfectly known. A crucial mechanism that generates strikes is that the outcome of a strike today can alter the relative advantages of the parties in future bargaining. In particular, a strike becomes more likely if one player stands to increase its future advantage substantially by winning a strike today. Collective bargaining institutions that condition the initial distribution (and potential redistribution) of such advantage can thus work to promote either industrial peace or strife. It will be seen that the latter is more likely when it is the institutionally favored party who stands to enlarge its advantage by winning a strike. Seen in this light, U.S. strike waves in the past century can be understood in part as a response to shifts in the institutional environment regulating labor relations.

The remainder of this chapter is organized as follows: Section 1.2 reviews existing attempts to explain strikes and clarifies the contribution of my model, Section 1.3 presents the model, Section 1.4 discusses the results in light of historical evidence, and Section 1.5 concludes.

## 1.2 Overview of Existing Models of Strikes

Strikes are challenging to theorize as actions by rational agents. The main difficulty lies in explaining why self-interested actors such as the union and management would resort to such a wasteful means of distributing the surplus from production, when both can be made better off by a peaceful agreement that allocates greater shares to each. Extant models of strikes (and bargaining more generally) have attempted to address this problem mostly by assuming some form of private information.

Perhaps the first such attempt was Ashenfelter and Johnson (1969). In their model, strikes occur as a mechanism to resolve information asymmetry between the union leadership and the rank and file regarding the profitability of the firm: when rank-and-file members demand a pay raise that the leadership privately knows is not winnable given the actual profitability of the firm, the leadership can nevertheless launch a strike to preserve its popularity, while the firm waits long enough for the hard realities of a strike to bring the expectations of lay union members in line with what the firm is willing to pay. There is a sense in which Ashenfelter and Johnson's account merely shifts the irrationality onto rank-and-file union members, but their work laid the basis for later one-sided asymmetric information models.

Thus, a class of 'screening' models have been subsequently developed (among others, by Sobel and Takahashi, 1983; Fudenberg and Tirole, 1983; Hayes, 1984; Tracy, 1987) which feature unions making a series of declining wage offers to the firm in an attempt to price-discriminate against more profitable firms. The idea is that more profitable firms would accept high wage offers early on rather than let profits dissipate from a protracted strike.

The screening models assume that offers are made at discrete intervals of time, for no apparent reason. This assumption lends these models an unappealing feature called the Coase Property (first conjectured by Coase, 1972, and later examined by Stokey, 1981 and Gul et al., 1986), which states that as the length of the interval between offers shrinks, wage offers as well as the duration of strike in equilibrium gets arbitrarily small. In other words, unless the union can commit to delays between offers, the firm ends up capturing all of the gains of trade, and there are essentially no strikes (See Kennan and Wilson, 1989 for an intuitive illustration of this result). The challenge, then, becomes one of explaining how the union can commit to delaying offers, or alternatively, why offers cannot be made continuously in practice (e.g. Hart, 1989 rationalizes delays in terms of transaction costs).

A closely related class of models overcomes this problem by allowing agents to choose the length of time between offers as a way of signalling their valuations. In these 'signalling' models (among others, by Admati and Perry, 1987; Cho, 1990; Cramton, 1992), a firm (union) can credibly signal its valuation (reservation wage) only by undertaking a costly delay until making its (first and final) offer. Once all private information is truthfully revealed through each party's timing of offers, the parties agree on the wage predicted by the Rubinstein (1982) model of bargaining with complete information. One problem with these models is that their predictions (using reasonable parameter values) appear to be grossly at odds with data on strike duration and settlement rates (Kennan and Wilson, 1989).

Whereas both the screening and signalling models identify the firm's unobserved profitability as the source of uncertainty generating disagreements and strikes, 'war of attrition' models locate the source of uncertainty in the parties' mutually unobserved costs of delaying agreement. Originally proposed by Smith (1974) to explore evolutionarily stable equilibria in animal behavior and later adapted to economic settings by Fudenberg and Tirole (1986) and Nalebuff and Riley (1985), these models assume a 'winner-take-all' setting in which the first party to back out of a contest loses all claim to the contested prize. Each contestant's optimal timing of capitulation depends on the probability distribution of both parties' cost per unit time of remaining in the contest (which is assumed to be common knowledge) and the privately known realizations of those costs.

One potential drawback of the models discussed thus far is that they rely crucially on the impossibility of conveying information except by going through a strike – an assumption that may not always be justified, especially if agents can learn from repeated interactions. Fernandez and Glazer (1991) demonstrated that such an assumption is unnecessary, by formulating a model of strikes under complete information. Although theoretically illuminating, their result relies on an implausibly complicated set of off-equilibrium strategies. A recent innovation by Schaller and Skaperdas (2020) obviates such complexity by supposing a preparatory stage preceding the outbreak of a strike (or some other form of contest such as a war or litigation), when up-front investments (arming, etc.) can be made to bolster one's chances of winning in the eventuality of a strike or, alternatively, one's bargained share if a strike can be avoided. Their finding is that a large enough asymmetry of power (in terms of cost of up-front investments and of effort expended in actual conflict) can make a strike ex-ante more profitable than bargaining for the more powerful player.

All of the models discussed above are formally dynamic in that they feature agents facing multiple (often infinite) periods in which they make choices (e.g. prepare for a strike; go on strike; make, withhold, reject, or accept offers) in anticipation of what happens in the next period. But they are also static in the sense that they treat collective bargaining and strikes essentially as a one-off affair, i.e. a game played only once when an existing contract expires, without regard to how the outcome of the current game might influence future negotiations and strikes. It would be reasonable, however, to suppose that forward-looking firms and unions interacting repeatedly over contract renewals would be mindful of how the outcome of the dispute in the current contract term might affect outcomes in future terms by way of setting a precedent. The widespread use of wage patterns in collective bargaining is indicative of the power of precedents (in the form of wages previously agreed on) to set the terms of future negotiations. But even more consequential may be the precedent of a strike won or lost, which could have a lasting impact on the morale of participants: for example, the trauma of defeat may weaken the capacity of a union to pull off a strike in the next contract and thereby undermine the credibility of its threat at the negotiating table.

This chapter proposes a model of strikes that is dynamic in this longer-term sense, in that it considers strikes and lockouts as a way of altering the bargaining power of the agents in future disputes. My aim is to show that when rational agents care enough about the future, strikes/lockouts can occur even if bargaining is Pareto-efficient in the stage game and there are no information asymmetries. This is an idea first formalized in the field of conflict economics by Garfinkel and Skaperdas (2000), who show that destructive conflict can occur in equilibrium if, for example, the victor gains permanent possession over the contested prize and is thereby freed from the imperative to arm itself every period in order to maintain its bargaining position under an 'Armed Peace' regime. Their fundamental insight is that when the future matters sufficiently, eliminating the cost of dealing with an adversary who would otherwise pose an ever-present threat may be well worth the immediate destruction brought on by open conflict.

The present chapter extends the approach by examining a case where conflict does not remove the fighting capacity of the losing side once and for all, but merely shifts the relative advantage of the contestants in future bargaining and fighting. Although similar in vein to McBride and Skaperdas (2007), who examine a case where one side can be eliminated from the contest by losing a series of battles with successively declining chances of winning, this study rules out such elimination and instead focuses on how the potential gains in future fighting capacity affects the decision to fight or bargain today, along a continuum of such potential gains. This feature makes the model more relevant to an industrial dispute setting, where losing a strike/lockout (even multiple times) doesn't usually mean that the management (if not the union) will be permanently incapacitated from fighting another day. Another novel feature of the model is that it highlights the role of institutional asymmetries that advantage one or the other player in generating conflict. Such asymmetries are likely present in any industrial relations regime that impose disparate rules of behavior on the parties of collective bargaining, and require special consideration in a model that pertains to strikes or other contests that are regulated by an external authority.

## 1.3 A Model of Strikes Under Complete Information

There are two players in the game, the union (denoted by $L$ for labor) and management (denoted $M$), who contest their shares over a pie either by fighting or by bargaining under the threat of fighting. I begin by describing what happens in the event of a fight (strike or lockout), since this eventuality constrains the choices of the parties in every contingency. Although strikes and lockouts may differ in their practical details, I treat them as equivalent in my model: they both entail a work stoppage whose sole function is to break the will of one side to resist and impose the other's will, in the event negotiations break down. The outcome of a strike/lockout is assumed to be binary and probabilistic: either $L$ or $M$ wins with a commonly known probability. I specify this probability as a contest success function (originally introduced by Tullock, 1980, later axiomized by Skaperdas, 1996, and extended by Clark and Riis, 1998 to allow for asymmetries), the workhorse model of conflict economics. Thus the probability that L wins a strike is given by

$$P(\text{L wins}) = P(\lambda(e_M, e_L)) = \frac{\lambda e_L}{\lambda e_L + (1-\lambda)e_M} \tag{1.1}$$

where $e_L$ and $e_M$ respectively denote effort levels devoted by each side to the contest. I assume effort to be costly and of a psychic nature: union officials have to exert mental energy in persuading members to join and stay in a picket line, management has to divert attention away from managing the business toward dissuading would-be strike participants, and so on. That effort mainly involves persuasion relates to a key aspect of a strike, namely that it is a form of collective action. Its success or failure depends on whether enough members can be persuaded to participate in the action and continue the disruption long enough to force management's capitulation.[1] But the same level of effort at persuasion can be more or less effective depending on the institutional environment. Thus when hiring

[1] Note that when efforts involve persuasion, the contest success function has a natural interpretation in Skaperdas and Vaidya (2012).

replacement workers during a strike is legal or widely accepted practice, it would be more challenging for the union to convince members that a strike will be effective, and conversely easier for management to convince workers to return to work.

The parameter $\lambda \in (0,1)$ can be thought of as capturing the effect of such legal or institutional setting on the ability of $L$ to hold together a strike. For simplicity I restrict $\lambda$ to take on only one of two discrete values: $\lambda \in \{\underline{\lambda}, \overline{\lambda}\}$, where $\lambda = \underline{\lambda}$ can be interpreted as a state where the union can only rely on the minimum protection offered by law for workers to collectively organize. Protections under the National Labor Relations Act such as sanctions against 'unfair labor practices' (i.e. the practice of penalizing employees for membership in a union) put a limit to what employers can do to undermine the power of employees to organize collectively, and thus make it feasible even for the most disadvantaged union to win a strike with non-zero probability.

$\overline{\lambda}$ can then be interpreted as the maximal state of power labor can enjoy under the existing institutional environment; from the perspective of management, $1-\overline{\lambda}$ represents the minimal institutional protection afforded to property rights. Labor laws in general have as their stated purpose the 'leveling of playing fields' for both labor and management, and proscribe the behavior of both parties in such a way that no one party can completely dominate the other. An example of a legal constraint on the labor side would be restrictions on the type of picket a union is allowed to organize. Under U.S. law, for example, if a picket is deemed so intimidating as to infringe upon the right of non-participants to continue working without restraint or coercion, the picket can be disbanded by a court injunction. The more restrictive such rules, the lower is $\overline{\lambda}$: since pickets of a given size would be rendered less effective in disrupting production, the union would need to work harder to persuade more co-workers to join the strike in order to achieve the same level of effectiveness.

Whether $\lambda$ takes on the lower or higher value depends on the outcome of the previous strike/lockout: a victorious strike for the union sets $\lambda = \overline{\lambda}$ until the next strike results in defeat for the union, at which point $\lambda$ is reset to $\underline{\lambda}$, and so on. In the meantime, $\lambda$ remains fixed in either the low state or high state when fighting does not occur, i.e. when $L$ and $M$ bargain and settle. The shifts to and from $\underline{\lambda}$ and $\overline{\lambda}$ may be due to explicit concessions made by the vanquished which directly affect the union's organizational capacities: employers may, for example, concede to a union-shop agreement (where it is legal) whereby new hires become union members by default. But perhaps more important is the effect that a victory or defeat has on the morale of the players. The experience of a victorious strike can go a long way toward convincing union members that building for a future strike is well worth the effort, and this increased confidence can be self-fulfilling especially when the success of collective action is predicated on every member's belief that every other member is confident enough not to desert their ranks and act as strikebreakers to save their own skins. In the Appendix I present a simple model that illustrates this collective action dynamic, which also points to some mechanisms by which $\lambda$ can shift as a result of a strike won or lost.

Henceforward I will refer to $\underline{\lambda}$ as the state in which $L$ is 'disadvantaged' ($M$ is 'advantaged') and $\overline{\lambda}$ as the state in which $L$ is 'advantaged' ($M$ is 'advantaged'). 'Advantage' here does not, strictly speaking, refer to power relative to the adversary: if $\underline{\lambda} > \frac{1}{2}$, for example, $L$ would be more powerful than $M$ even if it is described as 'disadvantaged'. Rather, 'advantage' denotes a player's position relative to the maximum power it can achieve by winning in a fight.

We are now ready for a formal description of the game. $M$ and $L$ play a stage game with the following sequence of events for two or more periods:

1. $M$ and $L$ seek to sign a new labor contract specifying a share of the unitary surplus (rent/profit) that goes to $L$ for that period (the last contract having just expired). Agreement on the terms of the contract can be reached either through bargaining (hereafter used interchangeably with 'negotiation' or 'settlement') or fighting ('strike/lockout' or 'open conflict'). The players observe the state $\lambda$ for that period

and form beliefs about whether fighting will occur; they then choose their effort levels $e = (e_L, e_M)$ to invest in anticipation of fighting or bargaining. Neither party observes the other's effort at this point.

2. The choices of effort levels are now revealed, and based on this information the players each simultaneously choose an enforcement strategy $\sigma \in \{F, B\}$, where $F$ denotes fighting and $B$ denotes bargaining. Let $S = (\sigma_L, \sigma_M)$ denote the profile of enforcement strategies. If $S = (F, F)$, $(F, B)$ or $(B, F)$ (i.e. if either player chooses to fight), then a strike/lockout occurs: a party choosing to fight effectively imposes the same choice on the other, although the party that did not anticipate fighting may have irreversibly chosen an effort level that is suboptimal. A negotiated settlement is reached only if $S = (B, B)$.

3. In the event of a strike, the surplus for that period shrinks by a factor $\varphi \in (0, 1)$, and the victor appropriates the entire surplus for that period. Next period's $\lambda$ is set to $\underline{\lambda}$ or $\overline{\lambda}$ depending on who emerges as the victor. If bargaining occurs, the parties split the unitary surplus for that period according to some predetermined share that is a function of the current-period state and effort levels. The current-period $\lambda$ then carries over to the next period unchanged.

It is worth noting at this point that even if open conflict can be avoided through a peaceful agreement, the outcome of this game will not be socially optimal: both sides would have expended some costly effort that contributes nothing to material production. A benevolent social planner would obviate such effort by imposing an arbitrary but incontestable share of the surplus to each side. The problem that arises in the absence of such a mythical creature is that efforts are often incontractible. In principal-agent game settings, for example, effort is deemed incontractible due to its limited observability; it is arguably even less observable in adversarial contests where maintaining secrecy is a paramount concern. As a result, the contestants in the game described above are incapable of committing to exert no effort either in the current period or in any subsequent period. This setup generates inefficiency in two ways: first, there is nothing to prevent the players from investing in socially wasteful effort as long as that investment is individually rational; second, as in Garfinkel and Skaperdas (2000), the fact that future exertion of effort can be reduced (or eliminated) by defeating the adversary in an open conflict can create an incentive to fight, even when bargaining would be the second-best outcome.

With this in mind, we will examine what happens when the stage game is played over two periods. In principle the number of periods could be extended to infinity, but we focus on a two-period setting both because it affords simplicity and because it permits a rigorous derivation of the rule of division under bargaining. In Section 1.3.1, we examine the outcome of a baseline one-shot game, which corresponds to the outcome in the final period; in Section 1.3.2 we move to the first period and look for conditions under which fighting can be an equilibrium; Section 1.3.3 explores conditions under which a bargaining equilibrium can hold.

### 1.3.1 Outcome of the stage game in period 2

Let $t = 1, 2$ denote the period. Solving backwards, we start from $t = 2$. $M, L$ are assumed to be risk-neutral maximizers of linear utility. Let $u_i^\sigma(\lambda, e); e = (e_i, e_j)$ denote player $i$'s one-period utility to choosing enforcement strategy $\sigma$ and effort $e_i$ given the state $\lambda$. Also let $P(\lambda, e)$ denote the probability $L$ wins a strike, as defined in (1.1). The players' one-period utility, in the event of fighting (or $\sigma_i = F$ for at least one player), are given as follows:

$$u_L^F(\lambda, e) = \varphi P(\lambda, e) - e_L = \varphi \cdot \frac{\lambda e_L}{\lambda e_L + (1-\lambda)e_M} - e_L,$$
$$u_M^F(\lambda, e) = \varphi(1 - P(\lambda, e)) - e_M = \varphi \cdot \frac{(1-\lambda)e_M}{\lambda e_L + (1-\lambda)e_M} - e_M.$$

where the first term in the right hand side of each expression represent the party's expected share of the surplus diminished by the strike.

$(u_L^F(\lambda, e), u_M^F(\lambda, e))$ constitutes a threat point that is inside the frontier of Pareto Efficient allocations obtainable when there is no fight. This means that any number of bargaining equilibria can be reached (and a fight avoided) as long as the allocation $(x, 1-x)$ of the unitary economic surplus under bargaining, where $x \in (0, 1)$ denotes the share going to L, has the feature that $u_L^F(\lambda, e) \leq x - e_L$ and $u_M^F(\lambda, e) \leq 1 - x - e_M$ To fix ideas, we will focus on a particular rule of allocation, often referred to as 'split the difference'. Letting $u_i^B(\lambda, e)$ denote the utility for player $i$ of bargaining, we shall derive an expression for $x$ such that:

$$u_M^B(\lambda, e) - u_M^F(\lambda, e) = u_L^B(\lambda, e) - u_L^F(\lambda, e).$$

More specifically:

$$1 - x - \varphi \cdot \frac{(1-\lambda)e_M}{\lambda e_L + (1-\lambda)e_M} - e_M = x - \varphi \cdot \frac{\lambda e_L}{\lambda e_L + (1-\lambda)e_M} - e_L$$

Solving for $x$, we obtain the following rule:

$$\begin{aligned} x &= \frac{1-\varphi}{2} + \varphi \cdot \frac{\lambda e_L}{\lambda e_L + (1-\lambda)e_M} - e_L, \\ 1 - x &= \frac{1-\varphi}{2} + \varphi \cdot \frac{(1-\lambda)e_M}{\lambda e_L + (1-\lambda)e_M} - e_M, \end{aligned}$$

where the first terms in the right hand sides correspond to equal splits in the efficiency gains from a strike averted, and the second terms represent the expected shares for each player under fighting (i.e. the threat point). Using this rule of allocation, we obtain the one-period payoffs under bargaining for each player:

$$
\begin{aligned}
u_L^B(\lambda, e) &= \frac{1-\varphi}{2} + \varphi \cdot \frac{\lambda e_L}{\lambda e_L + (1-\lambda)e_M} - e_L, \\
u_M^B(\lambda, e) &= \frac{1-\varphi}{2} + \varphi(1 - P(\lambda, e)) - e_M = \varphi \cdot \frac{(1-\lambda)e_M}{\lambda e_L + (1-\lambda)e_M} - e_M.
\end{aligned}
\tag{1.2}
$$

We now proceed to derive optimal efforts under fighting and bargaining, and then to derive indirect utilities under both scenarios. If both sides anticipate fighting, they will choose effort levels $e_i^* = e_i^*(\lambda, e_j)$, $i \neq j$ that satisfy the following first-order conditions:

$$
\begin{aligned}
\frac{\varphi\lambda(1-\lambda)e_M}{(\lambda e_L^* + (1-\lambda)e_M)^2} &= 1 \quad \text{L's FOC}, \\
\frac{\varphi\lambda(1-\lambda)e_L}{(\lambda e_L^{+}(1-\lambda)e_M^*)^2} &= 1 \quad \text{M's FOC},
\end{aligned}
\tag{1.3}
$$

where the left hand sides represent the marginal benefit of effort in terms of gains in expected share and the right hand sides the marginal cost. Solving this system of symmetric FOC's for optimal efforts under fighting, $e_i^* = e_i^F(\lambda)$, yields the following symmetric expressions:

$$
e_i^F(\lambda) = \varphi\lambda(1-\lambda), \; i = L, M. \tag{1.4}
$$

Note that effort intensity increases as the contest gets closer to an even fight, that is, as $\lambda$ approaches $\frac{1}{2}$. This result has some intuitive appeal. In sporting competitions at least, an informal but commonplace observation is that players do not exert as much effort when competing against a vastly superior or inferior team than they would if they were more evenly matched: when additional effort is not likely to change the outcome, trying too hard is simply not 'worth it'. An interesting implication is that there can be convex returns to being advantaged, and that 'leveling the playing field' is not necessarily welfare-enhancing from a societal perspective.

What about efforts under bargaining? Taking first-order conditions of (1.2) and solving for optimal bargaining efforts, $e_i^B(\lambda)$, we arrive at a solution identical to (1.4):

$$e_i^B(\lambda) = \varphi\lambda(1-\lambda),\ i = L, M. \tag{1.5}$$

In our particular setting, this result stems from a special feature of the 'split-the-difference' rule of allocation, namely that the marginal benefit of effort under bargaining is the same as under fighting. The implication is that at $t = 2$, both parties would choose the same levels of effort no matter what they anticipate, which effectively simplifies the game to the choice of $\sigma_i \in \{F, B\}$. Substituting efforts in (1.2) for (1.5), we obtain indirect utilities $u_i^B(\lambda, e) = v_i^B(\lambda)$ in the stage game when $S = (B, B)$ as follows:

$$\begin{aligned} v_L^B &= \frac{1-\varphi}{2} + \varphi\lambda^2, \\ v_M^B &= \frac{1-\varphi}{2} + \varphi(1-\lambda)^2. \end{aligned} \tag{1.6}$$

Note that the shares of the surplus going to each party under bargaining, which are simply the indirect utilities plus optimal efforts, are the following:

$$\begin{aligned} x &= \frac{1-\varphi}{2} + \varphi\lambda, \\ 1 - x &= \frac{1-\varphi}{2} + \varphi(1-\lambda). \end{aligned}$$

It is easily checked that the indirect utilities under fighting, when $\sigma_i = F$ for at least one player, are the following:

$$v_L^F(\lambda) = \varphi\lambda^2, \qquad v_M^F(\lambda) = \varphi(1-\lambda)^2.$$

Comparing the indirect utilities under fighting and bargaining, it is straightforward to see that choosing $\sigma_i = B$ is a weakly dominant strategy for both players. Weak dominance implies $S = (B, B)$ is not a unique Nash equilibrium of the stage game; nonetheless we shall assume that the players are able to coordinate on this equilibrium, and we can fix the players' equilibrium second-period payoffs, $v_{i2}(\lambda)$, as being identical to the expressions in (1.6).

### 1.3.2 Conditions for a fighting equilibrium

We now consider what happens in t = 1. Let $U_i^\sigma(\lambda, e)$ denote player i's expected two-period utility at the beginning of t = 1 as a function of enforcement strategy $\sigma$ and effort profile $e = (e_i, e_j)$ chosen at t = 1, as well as current-period state $\lambda$. Then if fighting occurs, we have that:

$$U_L^F(\lambda, e) = \varphi P(\lambda, e) + \delta\left(P(\lambda, e) \cdot v_{L2}(\overline{\lambda}) + (1 - P(\lambda, e)) \cdot v_{L2}(\underline{\lambda})\right) - e_L,$$
$$U_M^F(\lambda, e) = \varphi(1 - P(\lambda, e)) + \delta\left(P(\lambda, e) \cdot v_{M2}(\overline{\lambda}) + (1 - P(\lambda, e)) \cdot v_{M2}(\underline{\lambda})\right) - e_M,$$

where the second terms in the right hand sides capture the present values of expected second-period utilities, discounted by $\delta \in (0, 1)$. Note that $\delta$ doesn't have to be a time-invariant constant. It can be viewed as a function of the expected growth rate $g$ of the contestable surplus: e.g. $\delta = \beta(1 + g)$, where $\beta < 1$ represents the psychologically invariant time-discounting parameter. This implies that for high enough $g$, $\delta$ could exceed unity. However, I shall abide by the convention of bounding $\delta$ below 1: the qualitative results of the analysis below do not change by allowing $\delta > 1$, since agents in this model are not infinitely lived.

Similar to the stage game in the second period, we look for a rule of allocation under bargaining, $(y, 1 - y)$, where $y \in [0, 1]$ denotes the share of the surplus at $t = 1$ going to $L$ in the event of a negotiated settlement. Then the players' lifetime utilities under bargaining assume the following forms:

$$\begin{aligned} U_L^B(\lambda, e) &= y + \delta v_{L2}(\lambda) - e_L, \\ U_M^B(\lambda, e) &= 1 - y + \delta v_{M2}(\lambda) - e_M, \end{aligned} \tag{1.7}$$

where the second-period state $\lambda$ in the arguments of $v_{i2}(.)$ is the same as in the current period, because bargaining preserves the relative advantages of the players.

Again assuming the players 'split the difference', $y$ has to satisfy the condition:

$$U_L^B(\lambda, e) - U_L^F(\lambda, e) = U_M^B(\lambda, e) - U_M^F(\lambda, e).$$

Put differently:

$$\begin{aligned} & y + \delta v_{L2}(\lambda) - \varphi P(\lambda, e) - \delta(P(\lambda, e) \cdot v_{L2}(\overline{\lambda}) + (1 - P(\lambda, e)) \cdot v_{L2}(\underline{\lambda})) \\ & \quad = 1 - y + \delta v_{M2}(\lambda) - \varphi(1 - P(\lambda, e)) - \delta(P(\lambda, e) \cdot v_{M2}(\overline{\lambda}) + (1 - P(\lambda, e)) \cdot v_{M2}(\underline{\lambda})) \end{aligned}$$

After some algebra, it can be shown that:

$$y = y(\lambda, e) = \frac{1 - \varphi}{2} + \varphi \cdot \frac{\lambda(1 + \delta(\overline{\lambda} - \lambda))e_L - (1 - \lambda)\delta(\lambda - \underline{\lambda}e_M}{\lambda e_L + (1 - \lambda)e_M} \tag{1.8}$$

To facilitate interpretation, we plug in $\lambda = \underline{\lambda}, \overline{\lambda}$ to the argument of $y(., .)$ above:

$$\begin{aligned} y(\underline{\lambda}, e) &= \frac{1 - \varphi}{2} + \varphi(P(\underline{\lambda}, e) + \delta(\overline{\lambda} - \underline{\lambda})P(\underline{\lambda}, e)) \\ y(\overline{\lambda}, e) &= \frac{1 - \varphi}{2} + \varphi(P(\overline{\lambda}, e) - \delta\left(\overline{\lambda} - \underline{\lambda})(1 - P(\underline{\lambda}, e))\right) \end{aligned}$$

The first terms represent the same effciency gains from bargaining (evenly split between the players) as featured in (5). What is different is the threat point captured in the second term. $\varphi\delta(\overline{\lambda} - \underline{\lambda})$ is the present value of the difference in the second-period share going to $L$ that can result from winning versus losing a strike in the current period. When $\lambda = \underline{\lambda}$ in $t = 1$, $L$ has nothing to lose in $t = 2$ by fighting in $t = 1$: since $L$ is already disadvantaged, both bargaining today and losing a fight today will put it at exactly the same position tomorrow; but fighting gives $L$ a chance to win and become advantaged tomorrow, by a probability $P(\underline{\lambda}, e)$. Hence the possible present-discounted share gain $\varphi\delta(\overline{\lambda} - \underline{\lambda})$ weighted by the probability of winning, forms part of $L$'s threat point. When $\lambda = \overline{\lambda}$, on the other hand,

$L$'s share in $t = 2$ can only shrink as a result of fighting in $t = 1$, by a magnitude $\varphi\delta(\overline{\lambda} - \underline{\lambda})$ with probability $1 - P(\overline{\lambda}, e)$.

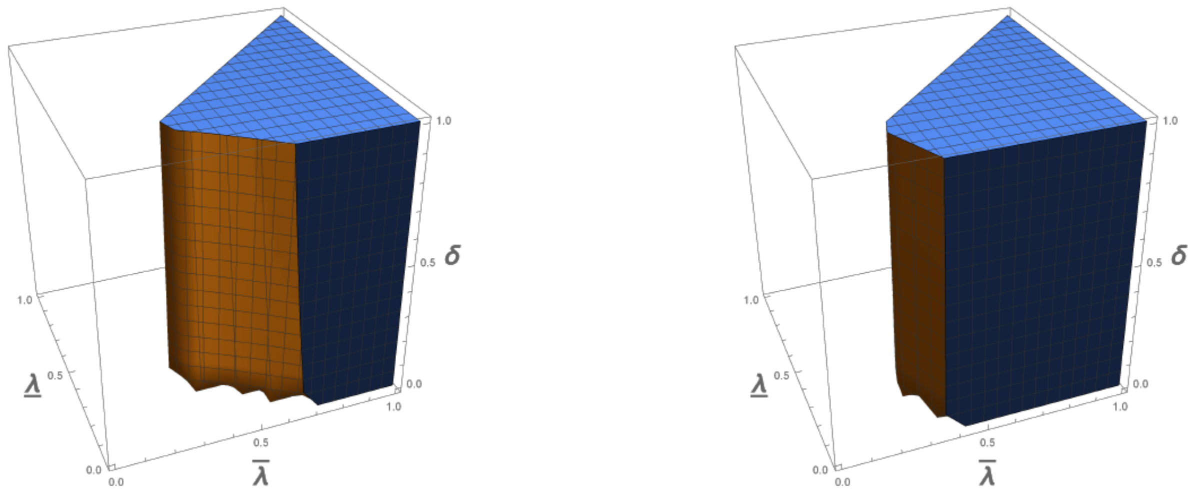
Plugging (1.8) into (1.7) yields lifetime utilities under bargaining as a function of the parameters and effort choices:

$$\begin{aligned} U_L^B(\lambda, e) &= \frac{\lambda(\frac{1+\varphi}{2} + \delta\varphi(\overline{\lambda} - \lambda + \lambda^2))e_L + (1-\lambda)(\frac{1-\varphi}{2} + \delta\varphi(\underline{\lambda} - \lambda + \lambda^2))e_M}{\lambda e_L + (1-\lambda)e_M} - e_L \\ U_M^B(\lambda, e) &= \frac{\lambda(\frac{1-\varphi}{2} - \delta\varphi(\overline{\lambda} - 2\lambda + \lambda^2))e_L + (1-\lambda)(\frac{1+\varphi}{2} - \delta\varphi(\underline{\lambda} - 2\lambda + \lambda^2))e_M}{\lambda e_L + (1-\lambda)e_M} - e_M \end{aligned} \tag{1.9}$$

We shall now compute optimal efforts under fighting and bargaining to then derive indirect lifetime utilities. Differentiating (13) with respect to efforts and setting equal to zero yields the following first-order conditions for optimal $t = 1$ fighting efforts $e_i^*$:

$$\begin{aligned} \frac{\varphi\lambda(1-\lambda)e_M(1 + \delta(\overline{\lambda}^2 - \underline{\lambda}^2))}{(\lambda e_L^* + (1-\lambda)e_M)^2} &= 1 \quad L\text{'s FOC} \\ \frac{\varphi\lambda(1-\lambda)e_L(1 + \delta((1-\underline{\lambda})^2 - (1-\overline{\lambda})^2))}{(\lambda e_L + (1-\lambda)e_M^*)^2} &= 1 \quad M\text{'s FOC} \end{aligned} \tag{1.10}$$

Comparing this set of FOC's with equivalent $t = 2$ FOC's in (1.3) is instructive. We see that the marginal benefit of effort is now rescaled by $1 + \delta(\overline{\lambda}^2 - \underline{\lambda}^2)$ for $L$ and by $1 + \delta((1-\underline{\lambda})^2 - (1-\overline{\lambda})^2)$ for $M$. Notice that $\delta(\overline{\lambda}^2 - \underline{\lambda}^2)$ (multiplied by $\varphi$ in the numerator) correspond to the present-discounted difference in $L$'s $t = 2$ utilities that can result from winning versus losing a strike in $t = 1$; $\delta((1-\underline{\lambda})^2 - (1-\overline{\lambda})^2))$ represents the analogous difference for $M$. Intuitively, the marginal benefits to effort are being inflated relative to those in a single period game because more is at stake than just the single-period surplus to be claimed by the winner: the contestants must also take into account the benefits that accrue to the winner in future periods as a result of becoming advantaged. This suggests, to a first approximation, that both sides would exert more effort toward fighting in $t = 1$ than in the final period. To verify this, we solve (1.10) for equilibrium fighting efforts $e_i^* = e_i^F(\lambda)$,




(a) $\lambda = \underline{\lambda}$ (b) $\lambda = \overline{\lambda}$

Figure 1.1: Range of parameter values for which $L$'s $t = 1$ fighting effort is higher than in $t = 2$

obtaining the following solutions:

$$
\begin{aligned}
e_L^F(\lambda) &= \frac{\varphi\lambda(1-\lambda)(1+\delta(\overline{\lambda}^2-\underline{\lambda}^2))^2\left[1+\delta((1-\underline{\lambda})^2-(1-\overline{\lambda})^2)\right]}{\left[1+\delta(1-\lambda)((1-\underline{\lambda})^2-(1-\overline{\lambda})^2)+\delta\lambda(\overline{\lambda}^2-\underline{\lambda}^2)\right]^2} \\
e_M^F(\lambda) &= \frac{\varphi\lambda(1-\lambda)(1+\delta(\overline{\lambda}^2-\underline{\lambda}^2))\left[1+\delta((1-\underline{\lambda})^2-(1-\overline{\lambda})^2)\right]^2}{\left[1+\delta(1-\lambda)((1-\underline{\lambda})^2-(1-\overline{\lambda})^2)+\delta\lambda(\overline{\lambda}^2-\underline{\lambda}^2)\right]^2}
\end{aligned}
\tag{1.11}
$$

We see that compared to $t = 2$ efforts (call them $e_{i,t=2}^F(\lambda)$), $t = 1$ fighting efforts ($e_{i,t=1}^F(\lambda)$) are scaled by a factor $C_i \equiv \frac{(1+\kappa_L)(1+\kappa_M)}{(1+(1-\lambda)\kappa_M+\lambda\kappa_L)^2} \times (1+\kappa_i)$, where $\kappa_L = \delta(\overline{\lambda}^2-\underline{\lambda}^2)$ and $\kappa_M = \delta((1-\underline{\lambda})^2-(1-\overline{\lambda})^2)$ respectively reflect what is at stake for $L$ and $M$, in terms of their discounted future returns to winning. It is unclear if $C_i$ is smaller or greater than 1. Figure 1.1 plots the range of values of the parameters $(\underline{\lambda}, \overline{\lambda}, \delta)$ for which $C_L > 1$, i.e. $e_{L,t=1}^F(\lambda) > e_{L,t=2}^F(\lambda)$, for states with $\lambda = \underline{\lambda}$ (panel (a)) and $\lambda = \overline{\lambda}$ (panel (b)).

What emerges from Figure 1.1 is that $L$ exerts less effort to fighting in $t = 1$ than in $t = 2$ if the institutional environment is too 'biased' against it, i.e. if both $\underline{\lambda}$ and $\overline{\lambda}$ are too low (formally, if $\overline{\lambda} + \underline{\lambda} < 1$). The intuition is as follows: when institutions are biased against $L$

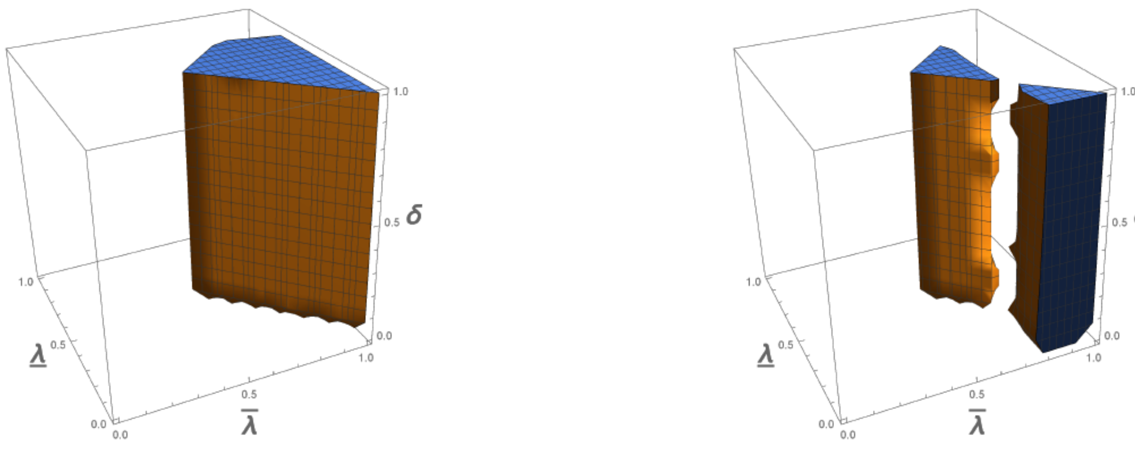
(henceforward I refer to $L$ as the 'underdog' in this situation), even if $L$ wins a fight in $t = 1$ it will have to exert more effort in $t = 2$ as a result of becoming advantaged (since $\overline{\lambda}$ would be closer to 1/2 than $\underline{\lambda}$ would be, and the contest would become more even); the reverse is true for the institutionally favored 'top dog' ($M$ in this case). This means the returns to winning in $t = 1$ (in terms of $t = 2$ utility) is lower for the underdog than for the top dog, which incentivizes the latter to exert more effort in $t = 1$ than it would in $t = 2$ under the same conditions. Combined with the already lopsided institutional environment, this additional effort on the part of the top dog has the effect of 'crowding out' the underdog's effort relative to that in $t = 2$: effort becomes too costly for the underdog. We next examine how $t = 1$ fighting efforts compare to $t = 1$ bargaining efforts. Differentiating (1.9) with respect to efforts yields the following first-order conditions:

$$\begin{aligned} \frac{\varphi\lambda(1-\lambda)e_M(1+\delta(\overline{\lambda}-\underline{\lambda}))}{(\lambda e_L^* + (1-\lambda)e_M)^2} &= 1 \quad L\text{'s FOC} \\ \frac{\varphi\lambda(1-\lambda)e_L(1+\delta(\overline{\lambda}-\underline{\lambda}))}{(\lambda e_L + (1-\lambda)e_M^*)^2} &= 1 \quad M\text{'s FOC} \end{aligned} \tag{1.12}$$

Comparing with the $t = 2$ FOC's in (1.3), we see that the marginal benefits are now scaled by $1 + \delta(\overline{\lambda} - \underline{\lambda})$, the second term (multiplied by $\varphi$ in the numerator) capturing the returns to winning in $t = 1$ in terms of $t = 2$ shares. Unlike in the FOC's for fighting effort, the returns to winning in terms of changes in $t = 2$ effort costs do not figure in the marginal benefits, since the rule of allocation $(y, 1-y)$ under bargaining are determined in such a way that cancels out the parties' identical $t = 2$ costs of effort. Solving (1.12) for $e_i^*$ yields the following symmetric expression for equilibrium effort choices under bargaining:

$$e_i^B(\lambda) = \varphi\lambda(1-\lambda)(1+\delta(\overline{\lambda}-\underline{\lambda})),\ i = L, M \tag{1.13}$$

It can be seen that bargaining effort in $t = 1$ is always larger than bargaining effort in $t = 2$. It is also apparent that unlike in $t = 2$, the level of effort now depends on whether it is geared toward fighting or bargaining. Which is larger? Figure 1.2 plots the range




(a) $\lambda = \underline{\lambda}$ (b) $\lambda = \overline{\lambda}$

Figure 1.2: Range of parameters for which $L$'s fighting effort is larger than bargaining effort, in $t = 1$

of parameters for which $e_L^B(\lambda) < e_L^F(\lambda)$ holds, for $\lambda = \underline{\lambda}$ (panel (a)) and $\lambda = \overline{\lambda}$ (panel (b)). Looking at panel (a) of Figure 1.2, one can see that fighting effort and bargaining effort are equal along the boundary that separates the space of $\underline{\lambda}$ and $\overline{\lambda}$ that corresponds to institutional bias in favor of $L$ from the $(\overline{\lambda}, \underline{\lambda})$ space corresponding to bias in favor of $M$. This is because in the absence of bias (i.e. when $\overline{\lambda} + \underline{\lambda} = 1$), the marginal returns to effort under bargaining and fighting become the same (FOC's (1.10) and (1.12)). When $L$ is the top dog in a position of disadvantage, it generally exerts more effort to fighting than to bargaining because the marginal returns to effort, in terms of gain in expected $t = 2$ utility, is higher under fighting. Panel (b) presents a more complicated picture. When $L$ is currently advantaged, it has to weigh the benefit of capturing the entire share of the pie today (minus the rent dissipation) against the potential loss in $t = 2$ utility from becoming disadvantaged, in deciding whether to fight or bargain. Conditional on fighting, its marginal return to effort is highest when either (i) it is the top dog enjoying only a slight bias in its favor and not too great an advantage, so that its chances of appropriating the entire share of the pie today is highly sensitive to its effort choice, or (ii) it is an underdog with large enough of an advantage and a small enough $\underline{\lambda}$ that it stands to lose a lot from losing a strike. Finally, we can obtain the value functions (or indirect lifetime payoffs) conditional on bargaining or fighting, $V_i^\sigma(\lambda)$, $i = L, M$; $\sigma = F, B$ by plugging in the optimal efforts $e_i^F(\lambda)$ and $e_i^B(\lambda)$ to the expressions for $U_i^F(\lambda, e)$ and $U_i^B(\lambda, e)$. (We omit the exact algebraic expressions for these value functions because they are unduly complicated and difficult to interpret).

We now consider two possible Nash equilibria in $t = 1$: a fighting equilibrium with $S = (F, F)$ and a bargaining equilibrium with $S = (B, B)$. Since effort levels now vary with the players' anticipated moves in the $t = 1$ stage game, ex-ante beliefs need to be explicitly considered. Let $\mu = (\mu_L, \mu_M)$ denote the profile of beliefs before effort choices are revealed, where player $i$'s belief $\mu_i = (\tilde{\sigma}, \hat{\sigma})$ consists of player $i$'s own belief $\tilde{\sigma} \in \{B, F\}$ about whether fighting will occur (e.g. $\tilde{\sigma} = F$ if player $i$ anticipates fighting), and player $i$'s belief about the other player's belief $\hat{\sigma} \in \{B, F\}$ about whether fighting will occur (e.g. $\hat{\sigma} = B$ if player $i$ believes player $j \neq i$ expects bargaining). As an example, $\mu = ((F, B), (B, B))$ indicates that while $M$ believes it is mutually understood that bargaining will occur and optimizes its effort toward bargaining as in (1.13), $L$ believes fighting will occur to $M$'s surprise – i.e. $L$ intends to ambush $M$. For the purpose of exploring possible equilibria, we rule out inconsistent beliefs such as $\mu = ((F, B), (F, B))$ which would imply each player mistakenly believes it can ambush the other player. Our focus is instead on two simple belief profiles, $\mu = ((B, B), (B, B))$ and $\mu = ((F, F), (F, F))$, and examine under what conditions these beliefs would be consistent with actual play. Without loss of generality, we assume $\lambda = \underline{\lambda}$ in $t = 1$, i.e. $L$ begins from a position of disadvantage. Since the players are symmetric, the results that follow would apply to whichever side that begins from a disadvantaged position. We begin with $\mu = ((F, F), (F, F))$. If belief in the inevitability of fighting is mutually shared, $L$ and $M$ will choose effort levels $e_i^F(\lambda)$. But having invested $e_i^F(\lambda)$, could players decide ex-post to not fight after all? If, upon observing each other's effort, the parties conclude that playing $S = (B; B)$ would make both strictly better off, then playing $\sigma = B$, $i = L, M$ becomes a weakly dominant strategy. This would be the case if the following condition holds:

$$V_i^F(\underline{\lambda}) \leq U_i^B(\underline{\lambda}, e^F(\underline{\lambda}))\,\forall i, \tag{1.14}$$

subject to the feasibility constraint:

$$0 \leq y(\underline{\lambda}, e^F(\underline{\lambda})) \leq 1. \tag{1.15}$$

Condition (1.15) merely states that any allocation to $L$ under bargaining can neither be negative nor exceed the surplus to be produced in that period. A stronger, but arguably more intuitive, condition may be that $L$'s share, when $L$ is the disadvantaged player, cannot exceed its hypothetical second-period share $x$ if it becomes advantaged, i.e. if $\lambda = \overline{\lambda}$ in $t = 2$. Thus:

$$y(\underline{\lambda}, e^F(\underline{\lambda})) < x(\overline{\lambda}) = \frac{1-\varphi}{2} + \varphi\overline{\lambda},$$

where the right side of the inequality represents $L$'s $t = 2$ indirect utility plus its effort (which equates to share of the surplus). To see if $L$ has an incentive to deviate from initial beliefs and choose $\sigma = B$ under certain conditions, I present in Figure 1.3 a 3-dimensional contour plot (left) of the space of $\overline{\lambda}, \underline{\lambda}$ and $\delta$ that satisfies (1.14) subject to (1.15), setting $\varphi = 0.9$. The solid area represents the set of parameter values that induces $L$ to choose $\sigma = B$ ex-post; the complement of this set represents parameter values that induces $L$ to stick to fighting. An alternative representation of this space is given in the right panel, which is a two-dimensional contour map of the values of $\delta$ that makes $L$ indifferent between fighting and bargaining. The oval-shaped hull in the front face of the solid area in the left panel (or the semi-crater in the northern region of the contour map in the right panel) thus represents the space of the parameters for which $L$ sticks to fighting. Let $\Gamma_i = \{(\underline{\lambda}, \overline{\lambda}, \delta) | U_i^B(\underline{\lambda}, e^F(\underline{\lambda})) < 0 \cup y(\underline{\lambda}, e^F(\underline{\lambda})) < 0 \cup \frac{1-\varphi}{2} + \varphi\overline{\lambda} < y(\underline{\lambda}, e^F(\underline{\lambda}))\}$, $i = L, M$

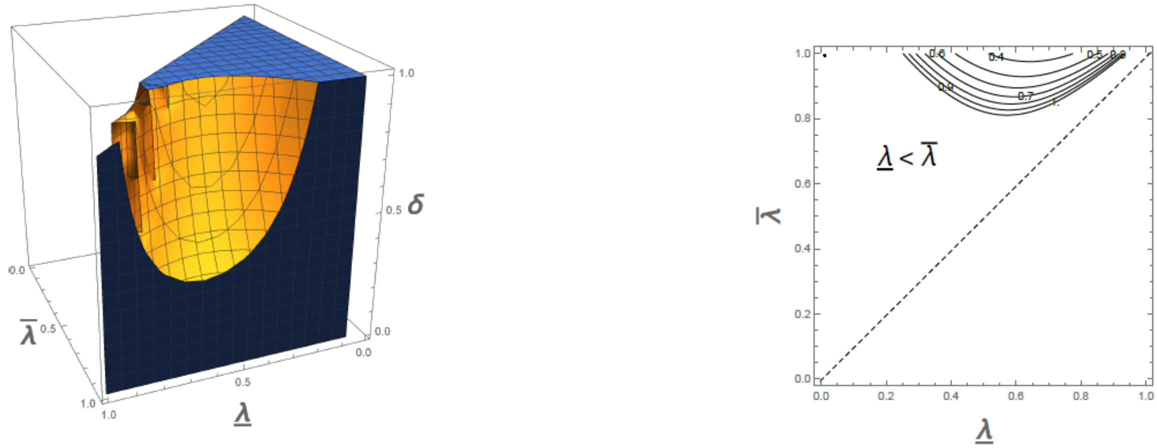



Figure 1.3: Range of parameters that induces $L$ to choose bargaining ex-post, conditional on both having chosen fighting effort

denote this space. That $\Gamma_i$ is a non-empty set for at least one $i$ is sufficient to show the existence of a fighting equilibrium: whenever $(\underline{\lambda}, \overline{\lambda}, \delta) \in \Gamma_i$, player $i$ fights given that efforts $e^F(\underline{\lambda})$ have been invested; given that fighting is mutually anticipated, efforts indeed equal $e^F(\underline{\lambda})$; thus the beliefs $\mu = ((F, F), (F, F))$ become self-fulfilling. Interestingly, the set of parameter values that induces $M$ to stick to fighting, $\Gamma_M$, turns out to be identical to $\Gamma_L$. We can see this by comparing the condition $V_i^F(\underline{\lambda}) > U_i^B(\underline{\lambda}, e^F(\underline{\lambda}))$ for $i = L$ with that for $i = M$. Letting $P(\underline{\lambda}, e^F(\underline{\lambda})) \equiv \underline{P}$ to economize on notation, we start with $L$:

$$
\begin{aligned}
&V_L^F(\underline{\lambda}) > U_L^B(\underline{\lambda}, e^F(\underline{\lambda})) \\
&\iff \varphi \underline{P} + \delta(\underline{P} v_{L2}(\overline{\lambda}) + (1 - \underline{P}) v_{L2}(\underline{\lambda})) - e_L^F(\underline{\lambda}) \\
&\qquad\qquad > \frac{1-\varphi}{2} + \varphi \underline{P}(1 + \delta(\overline{\lambda} - \underline{\lambda})) + \delta v_{L2}(\underline{\lambda}) - e_L^F(\underline{\lambda}) \\
&\iff \delta \underline{P}(v_{L2}(\overline{\lambda} - v_{L2}(\underline{\lambda})) - \varphi \underline{P} \delta(\overline{\lambda} - \underline{\lambda}) = \varphi \delta \underline{P}(\overline{\lambda}^2 - \underline{\lambda}^2 - (\overline{\lambda} - \underline{\lambda})) > \frac{1-\varphi}{2} \\
&\qquad\qquad \iff \varphi \delta \underline{P}(\overline{\lambda} - \underline{\lambda})(\overline{\lambda} + \underline{\lambda} - 1) > \frac{1-\varphi}{2}. \qquad (1.16)
\end{aligned}
$$

The equivalent condition for $M$ is:

$$\begin{aligned}
&V_L^F(\underline{\lambda}) > U_L^B(\underline{\lambda}, e^F(\underline{\lambda})) \\
&\iff \varphi(1-\underline{P}) + \delta(\underline{P}v_{M2}(\overline{\lambda}) + (1-\underline{P})v_{M2}(\underline{\lambda})) - e_M^F(\underline{\lambda}) \\
&\qquad\qquad > \frac{1+\varphi}{2} - \varphi\underline{P}(1+\delta(\overline{\lambda}-\underline{\lambda})) + \delta v_{M2}(\underline{\lambda}) - e_M^F(\underline{\lambda}) \\
&\iff \delta\underline{P}(v_{M2}(\overline{\lambda} - v_{M2}(\underline{\lambda})) + \varphi\underline{P}\delta(\overline{\lambda}-\underline{\lambda}) \\
&\qquad\qquad = \varphi\delta\underline{P}((1-\overline{\lambda})^2 - (1-\underline{\lambda})^2 + (\overline{\lambda}-\underline{\lambda})) > \frac{1-\varphi}{2} \\
&\qquad\qquad \iff \varphi\delta\underline{P}(\overline{\lambda}-\underline{\lambda})(\overline{\lambda}+\underline{\lambda}-1) > \frac{1-\varphi}{2}. \qquad (1.17)
\end{aligned}$$

To make sense of this identity, notice that the left hand side of the inequality in the second-to-last line in (1.16) represents $L$'s expected returns to winning in terms of $t = 2$ utility, $\delta\underline{P}(v_{L2}(\overline{\lambda}) - v_{L2}(\underline{\lambda}))$, minus that part of $L$'s share under bargaining that compensates $L$ for its expected $t = 2$ share gain from fighting, $\varphi\underline{P}\delta(\overline{\lambda}-\underline{\lambda})$, which $L$ would have to forego by choosing to fight ($L$'s expected share under fighting in $t = 1$ and the part of $L$'s share under bargaining, $y$, that compensates $L$ for this same expected share are both equal to $\varphi\underline{P}$, and therefore cancel out). For fighting to be optimal, this net gain has to outweigh the efficiency loss from fighting, represented by $\frac{1-\varphi}{2}$ in the right hand side. For $M$, the net gain (before efficiency losses) from fighting is its expected utility loss in $t = 2$ as a result of becoming disadvantaged, $\delta\underline{P}(v_{M2}(\overline{\lambda}) - v_{L2}(\underline{\lambda}))$, plus the benefit of not having to compensate $L$ for its $t = 2$ expected share gain, $\varphi\underline{P}\delta(\overline{\lambda}-\underline{\lambda})$, in the event of bargaining; again, this net gain has to outweigh the efficiency loss from fighting. It is no accident that the net gain of fighting (before efficiency losses) turns out be identical for both parties: the split-the-difference rule of allocation under bargaining, by construction, sets these quantities equal for the two sides.

Put another way, if the potential reward to becoming advantaged in $t = 2$ is big enough for the disadvantaged player in $t = 1$ to justify its immediate losses incurred by fighting (in terms of rent dissipation and forfeiture of a certain share it could have secured under bargaining), then the potential cost of becoming disadvantaged in $t = 2$ for the currently advantaged player is necessarily small enough, compared to the price of paying for peace, to justify fighting and the rent dissipation that it entails. And this has to do with the convex (concave) returns to winning for the top (under-) dog alluded to earlier. The top dog ($L$) who finds itself at a disadvantage ($\lambda = \underline{\lambda}$) in $t = 1$ not only stands to gain a higher share in $t = 2$ if it wins a strike in $t = 1$, but also to expend less effort in $t = 2$ than if it had lost the strike; the underdog ($M$), by contrast, stands to lose the same $t = 2$ share but also to expend less effort in t = 2 (since efforts are symmetric in $t = 2$), if it loses in $t = 1$. This means that conditional on efforts $e^F_{i,t=1}(\lambda)$ already invested, fighting can be more efficient than bargaining: it gives both sides an opportunity to reduce $t = 2$ efforts if the top dog wins in $t = 1$. Bargaining, on the other hand, ensures that the top dog will remain disadvantaged in the next period, forcing both sides to expend more effort. Graphically, the threat points now lie outside the Pareto frontier of allocations feasible under bargaining, and splitting the difference makes each party worse off relative to their threat points, by exactly the same magnitudes.

Note that the foregoing conditional efficiency argument only holds if it is the top dog that is at a disadvantage in $t = 1$: a necessary condition for (1.16) and (1.17) to hold is $\overline{\lambda}+\underline{\lambda}-1 > 0$ when $L$ is disadvantaged (when $M$ is disadvantaged, the equivalent condition is $\overline{\lambda}+\underline{\lambda}-1 < 0$). The flip side of this argument is that the underdog has to occupy a position of advantage in $t = 1$: Thus a strike cannot be an equilibrium outcome if either bias does not exist or if advantage lies with the side for whom institutions are biased in favor of. It is as if fighting functions as a mechanism for correcting an incongruence between advantage and institutional bias, as if the purpose of fighting is to restore advantage to its 'rightful' owner, the top dog.

Further examination of (1.16) yields some additional insights. The left hand side of the inequality is increasing in $\varphi$ (it can be checked that $\underline{P}$ is not a function of $\varphi$) while the reverse is true for the right hand side, which means $\Gamma_i$ expands as $\varphi \to 1$: the less destructive are strikes, the more likely their occurrence. Also, if $\overline{\lambda}+\underline{\lambda}-1$ (call it 'bias') is held fixed, the left hand side is increasing, to a first approximation, in $\overline{\lambda}-\underline{\lambda}$ (call it the 'stakes' of a fight) and in $\delta$: the more the future matters and the greater the difference winning can make, the more likely is fighting to occur. Similarly, holding stakes fixed, greater bias in favor of the disadvantaged player makes fighting more likely. The exact effect of bias ($B \equiv \overline{\lambda}+\underline{\lambda}-1$) and the stakes ($D \equiv \overline{\lambda}-\underline{\lambda}$) on the left hand side of (1.16) is complicated by the fact that $\underline{P}$ is also a function of those terms:

$$\underline{P} = P(\underline{\lambda}, e^F(\underline{\lambda})) = \frac{\underline{\lambda}(1+\delta(\overline{\lambda}+\underline{\lambda})(\overline{\lambda}-\underline{\lambda}))}{1+\delta(\overline{\lambda}-\underline{\lambda})(1+(\overline{\lambda}+\underline{\lambda}-1)(2\underline{\lambda}-1))}. \tag{1.18}$$

Nonetheless it can be seen that $\partial\underline{P}/\partial B$ and $\partial\underline{P}/\partial D$ will only be of second-order importance relative to the direct effect of $D$ and $B$ on the left hand side of (1.16); both $B$ and $D$ appear in both the numerator and denominator of (1.18), with the same sign except for $B$ if $\underline{\lambda} < 1/2$ (in which case $\partial\underline{P}/\partial B$ ¿ 0, and the net effect of an increase in $B$ on the left hand side of (1.16) will be amplified). A formal comparative static analysis in Appendix A.2 shows that the left hand side of (1.16) is always increasing in the magnitude of bias and nearly always increasing in the stakes of fighting.

We now condense the results of the analysis regarding a fighting equilibrium into the following proposition:

**Proposition 1.1.** *a fighting equilibrium exists if (i) institutions are biased ($\overline{\lambda}+\underline{\lambda} \neq 1$) in favor of the player who finds itself at a disadvantage in $t=1$, and (ii) the bias $B$ ($=\overline{\lambda}+\underline{\lambda}-1$ if $L$ is at a disadvantage, $=1-\overline{\lambda}-\underline{\lambda}$ if $M$ is at a disadvantage), difference in future advantage that is at stake $D \equiv \overline{\lambda}-\underline{\lambda}$, the importance of the future $\delta$, and the surplus shrinkage factor $\varphi$ are jointly large enough.*

### 1.3.3 Conditions for a bargaining equilibrium

We now turn to the viability of a bargaining equilibrium, supported by $\mu = ((B,B),(B,B))$. For $S = (B,B)$ to be an equilibrium, two conditions have to be met: first, neither player must have an ex-ante incentive for unilateral deviation, holding the other player's beliefs fixed. That is to say, the following condition must hold:

$$V_i^B(\lambda) \geq U_i^F(\lambda, e_i^{DF}(\lambda), e_{j\neq i}^B(\lambda)),\ \forall i, \tag{1.19}$$

where $e_i^{DF}$ denotes player $i$'s effort optimized toward ambushing player $j$, who believes bargaining will occur and invests the equilibrium bargaining effort. Second, even if both players invest $e_i^B(\lambda)$ as in (1.11), either player may still have an ex-post incentive to choose $\sigma_i = F$, and this must be precluded:

$$V_i^B(\lambda) + e_i^B(\lambda) \geq U_i^F(\lambda, e^B(\lambda)) + e_i^B(\lambda),\ \forall i. \tag{1.20}$$

Again, without loss of generality, we assume $L$ starts from a position of disadvantage ($\lambda = \underline{\lambda}$ in $t = 1$). We shall first derive $e_L^{DF}(\underline{\lambda})$ by taking the first-order condition for $U_L^F(\underline{\lambda}, e_M^B(\underline{\lambda}), e_L)$ with respect to $e_L$, after substituting $e_M^B(\underline{\lambda}) = \varphi\underline{\lambda}(1-\underline{\lambda})(1+\delta(\overline{\lambda}-\underline{\lambda}))$:

$$\frac{\varphi^2(1-\underline{\lambda})^2(1+\delta(\overline{\lambda}-\underline{\lambda}))(1+\delta(\overline{\lambda}^2-\underline{\lambda}^2))}{\left[\varphi(1-\underline{\lambda})^2(1+\delta(\overline{\lambda}-\underline{\lambda}))-e_L^*\right]^2} \leq 1.$$

Solving for $e_L^*$ yields:

$$e_L^* = e_L^{DF}(\underline{\lambda}) = \varphi(1-\underline{\lambda})(1+\delta(\overline{\lambda}-\underline{\lambda}))\left[\underline{\lambda} + \sqrt{\frac{1+\delta(\overline{\lambda}-\underline{\lambda})(\overline{\lambda}+\underline{\lambda})}{1+\delta(\overline{\lambda}-\underline{\lambda})}} - 1\right] \geq 0$$

Note that a corner solution ($e_L^* = 0$) necessarily implies deviation is unprofitable for $L$. Below we proceed on the assumption that an interior solution exists. Comparing the above expression for $e_L^{DF}$ with that for $e_L^B(\underline{\lambda})$ in (1.13), we see that $L$'s deviation effort will be larger than bargaining effort whenever $L$ enjoys institutional bias in its favor, i.e. when $\overline{\lambda} + \underline{\lambda} > 1$ and thus $\sqrt{\frac{1+\delta(\overline{\lambda}-\underline{\lambda})(\overline{\lambda}+\underline{\lambda})}{1+\delta(\overline{\lambda}-\underline{\lambda})}} - 1 > 0$. This is intuitive: if a disadvantaged top dog deviates to fighting, it would do so because it has more to gain (in terms of $t = 2$ expected utility) by disrupting the status quo and attempting to reclaim the advantage that is 'rightfully' due to it, than by perpetuating its own disadvantage; therefore it would devote greater effort to fighting than to bargaining. After making the substitutions $i = L$, $\lambda = \underline{\lambda}$, $e_M^B(\underline{\lambda}) = \varphi\underline{\lambda}(1-\underline{\lambda})(1+\delta(\overline{\lambda}-\underline{\lambda}))$, and $e_L^{DF}(\underline{\lambda}) = \varphi(1-\underline{\lambda})(1+\delta(\overline{\lambda}-\underline{\lambda}))\left[\underline{\lambda} + \sqrt{\frac{1+\delta(\overline{\lambda}-\underline{\lambda})(\overline{\lambda}+\underline{\lambda})}{1+\delta(\overline{\lambda}-\underline{\lambda})}} - 1\right]$, the condition (1.19) becomes:

$$
\begin{aligned}
&V_L^B(\underline{\lambda}) \geq U_L^F(\underline{\lambda}, e_L^{DF}(\underline{\lambda}), e_M^B(\underline{\lambda})) \\
&\iff \varphi\delta(\overline{\lambda}-\underline{\lambda})(1+\overline{\lambda}-\underline{\lambda}) \leq \frac{1-\varphi}{2} + 2\varphi(1-\underline{\lambda}) \\
&\qquad \times \left(\sqrt{1+\delta(\overline{\lambda}-\underline{\lambda}))(1+\delta(\overline{\lambda}-\underline{\lambda})(\overline{\lambda}+\underline{\lambda}))} - 1\right) \equiv RHS. \quad (1.21)
\end{aligned}
$$

Suppose $L$ is the top dog, i.e. $\overline{\lambda} + \underline{\lambda} > 1$. Then $RHS$ in the above inequality is bounded above and below in the following way:

$$\frac{1-\varphi}{2} + 2\varphi(1-\underline{\lambda})\delta(\overline{\lambda}-\underline{\lambda}) < RHS < \frac{1-\varphi}{2} + 2\varphi(1-\underline{\lambda})\delta(\overline{\lambda}-\underline{\lambda})(\overline{\lambda}-\underline{\lambda}).$$

Now suppose $L$ is the underdog, i.e. $\overline{\lambda} + \underline{\lambda} < 1$. Then the bounds are reversed:

$$\frac{1-\varphi}{2} + 2\varphi(1-\underline{\lambda})\delta(\overline{\lambda}-\underline{\lambda})(\overline{\lambda}-\underline{\lambda}) < RHS < \frac{1-\varphi}{2} + 2\varphi(1-\underline{\lambda})\delta(\overline{\lambda}-\underline{\lambda}).$$

Thus if $L$ is the top dog, a necessary condition for deviation to be ex-ante unprofitable for $L$ is:

$$\varphi\delta(\overline{\lambda}-\underline{\lambda})(1+\overline{\lambda}-\underline{\lambda}) < \frac{1-\varphi}{2} + 2\varphi(1-\underline{\lambda})\delta(\overline{\lambda}-\underline{\lambda})(\overline{\lambda}+\underline{\lambda})$$
$$\Longleftrightarrow \varphi\delta(\overline{\lambda}-\underline{\lambda})(2\underline{\lambda}-1)(\overline{\lambda}+\underline{\lambda}-1) < \frac{1-\varphi}{2}. \quad (1.22)$$

One can see that this condition will be violated if $\varphi$, $\delta$, $\overline{\lambda}-\underline{\lambda}$, and $\overline{\lambda}+\underline{\lambda}$ are jointly high enough, unless $\underline{\lambda} < 1/2$. A sufficient condition that precludes deviation ex-ante is:

$$\varphi\delta(\overline{\lambda}-\underline{\lambda})(1+\overline{\lambda}-\underline{\lambda}) < \frac{1-\varphi}{2} + 2\varphi(1-\underline{\lambda})\delta(\overline{\lambda}-\underline{\lambda})$$
$$\Longleftrightarrow \varphi\delta(\overline{\lambda}-\underline{\lambda})(\overline{\lambda}+\underline{\lambda}-1) < \frac{1-\varphi}{2}. \quad (1.23)$$

If $L$ is the underdog, the necessary condition for excluding ex-ante deviation (corresponding to (1.22)) turns out to be identical to (1.23), while the corresponding sufficient condition is identical to (1.22). While the necessary condition will always hold since $\overline{\lambda}+\underline{\lambda}-1$ in the left hand side of (1.23) will be negative in this case, the sufficient condition can be violated if $\underline{\lambda}$ is small enough relative to 1/2 and $\overline{\lambda}+\underline{\lambda}-1$ is negative enough (given, of course, high enough values of $\varphi$, $\delta$, and $\overline{\lambda}-\underline{\lambda}$). However, a corner solution becomes likely when $\underline{\lambda}$ and $\overline{\lambda}+\underline{\lambda}-1$ are too small, which suggests the sufficient condition for precluding $L$'s deviation would likely hold when we assume an interior solution.

None of these statements are enough to precisely characterize the conditions under which deviation is or is not profitable for $L$, except in the case where $L$ is the top dog and $\underline{\lambda} > 1/2$. So we rely on visual aid. Plotting the range of parameter values (again, setting $\varphi = 0.9$) that satisfy (1.21) yields Figure 1.4. Again, the oval-shaped hull in the front-facing plane of the 3-dimensional plot on the left panel (or the semi-crater in the northern region of the 2-dimensional contour map on the right panel) represent the space of parameter values where the inequality (1.21) is reversed, so that $L$ has an ex-ante incentive to deviate. The striking

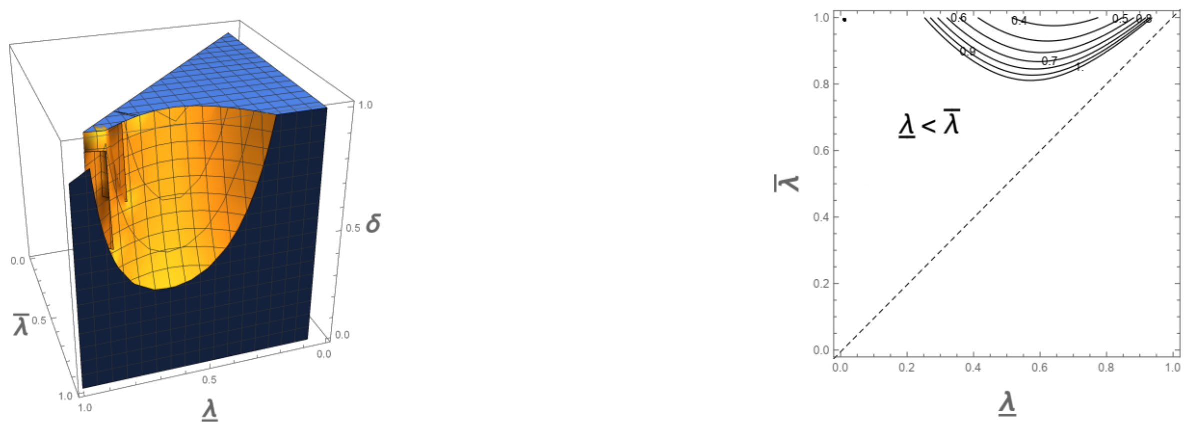



Figure 1.4: Range of parameters that induces $L$ to choose bargaining ex-ante

similarity with Figure 1.3 is explained by the fact that conditions (1.22) and (1.23) closely resemble the complement of the condition (1.16). The qualitative dynamic underlying the two Figures is the same: a disadvantaged top dog has an incentive to disrupt the status quo rather than sustain it. We now examine the condition ((1.19)) that would preclude $M$'s deviation ex-ante, when $\lambda = \underline{\lambda}$. Using the same procedure as before, we can derive $M$'s deviation effort:

$$e_M^{DF}(\underline{\lambda}) = \max\left\{ \varphi\underline{\lambda}(1+\delta(\overline{\lambda}-\underline{\lambda}) \left[\sqrt{\frac{1+\delta(\overline{\lambda}-\underline{\lambda})(2-\overline{\lambda}-\underline{\lambda})}{1+\delta(\overline{\lambda}-\underline{\lambda})}} - \underline{\lambda}\right], 0 \right\}.$$

Comparing with $e_M^B(\underline{\lambda}) = \varphi\underline{\lambda}(1+\delta(\overline{\lambda}-\underline{\lambda}))$, we see that $e_M^{DF}(\underline{\lambda}) > e_M^B(\underline{\lambda})$ if $2-\overline{\lambda}-\underline{\lambda} > 1$, or $\overline{\lambda}+\underline{\lambda} < 1$, i.e. if $M$ is the top dog. After making the necessary substitutions, condition (1.19) for $M$ becomes:

$$\begin{aligned} & V_M^B(\underline{\lambda}) \geq U_M^F(\underline{\lambda}, e_L^B(\underline{\lambda}), e_M^{DF}(\underline{\lambda})) \\ & \iff 2\varphi(1+\delta(\overline{\lambda}-\underline{\lambda}))\left[1-\sqrt{\frac{1+\delta(\overline{\lambda}-\underline{\lambda})(2-\overline{\lambda}-\underline{\lambda})}{1+\delta(\overline{\lambda}-\underline{\lambda})}}\right] \leq \frac{1-\varphi}{2}. \quad (1.24) \end{aligned}$$

It is immediately seen that (1.24) always holds if $2-\overline{\lambda}-\underline{\lambda} > 1$: when $M$ is the top dog who also currently occupies a position of advantage, it has no incentive to deviate to fighting and thereby risk ceding the advantage in the future. If $M$ is the underdog, on the other hand, (1.24) can be violated if $\varphi$, $\delta$, $\overline{\lambda} - \underline{\lambda}$ and $\overline{\lambda} + \underline{\lambda}$ are jointly high enough. The intuition is fundamentally the same as in the fighting equilibrium discussed above: an underdog who is currently advantaged is tempted to fight because the opportunity to appropriate the entire share of the contestable surplus today outweighs the potential utility loss tomorrow, which will be mitigated by the reduction in effort enjoyed by both parties in the event of the underdog's defeat. Thus it is no coincidence that the contour plot of parameter values satisfying (1.24) looks much the same as Figure 1.4 (not shown).

Finally, we check condition (1.20). Because $e_L^B(\lambda) = e_M^B(\lambda)$, we have the convenient result that $P(\lambda, e^B(\lambda)) = \lambda$. Thus for $L$, condition (1.20) translates into the following:

$$
\begin{aligned}
&V_L^B(\underline{\lambda}) + e_L^B(\underline{\lambda}) = y(\underline{\lambda}, e) + \delta v_{L2}(\underline{\lambda}) \\
&= \frac{1-\varphi}{2} + \varphi(\underline{\lambda} + \delta(\overline{\lambda} - \underline{\lambda})\underline{\lambda}) + \delta(\frac{1-\varphi}{2} + \varphi\underline{\lambda}^2) \\
&\geq V_L^F(\underline{\lambda}) + e_L^B(\underline{\lambda}) = \varphi\underline{\lambda} + \delta(\underline{\lambda} v_{L2}(\overline{\lambda}) + (1-\underline{\lambda}) v_{L2}(\underline{\lambda})) \\
&= \varphi\underline{\lambda} + \delta(\frac{1-\varphi}{2} + \varphi\underline{\lambda}^2 + \underline{\lambda}\varphi(\overline{\lambda}^2 - \underline{\lambda}^2)) \\
&\qquad \iff \delta\varphi\underline{\lambda}(\underline{\lambda}(1-\underline{\lambda}) - \overline{\lambda}(1-\overline{\lambda})) \leq \frac{1-\varphi}{2}. \quad (1.25)
\end{aligned}
$$

Since $\lambda(1-\lambda)$ is maximized at $\lambda = 1/2$, we can see that $\underline{\lambda}(1-\underline{\lambda}) - \overline{\lambda}(1-\overline{\lambda})$ will be positive if $\underline{\lambda}$ is closer to $1/2$ than $\overline{\lambda}$ is, i.e. when $L$ is the top dog. Thus (1.25) says $L$ would have an ex-post incentive to deviate to fighting if it is the top dog and $\varphi$, $\delta$, and $\underline{\lambda}$ are jointly high enough to violate the inequality. It can be shown that condition (1.20) for $M$ is identical to (1.25), the interpretation being that $M$ is tempted to deviate when it is the underdog currently enjoying advantage. The plot of parameter values satisfying (1.25) is visually indiscernible from Figure 1.4 (not shown, to avoid repetition) suggesting that the conditions that trigger ex-ante deviation and ex-post deviation are practically the same. Nonetheless, we can still make a statement regarding which condition for a bargaining equilibrium is more restrictive. Since $U_i^F(\lambda, e^B(\lambda)) < U_i^F(\lambda, e_i^{DF}(\lambda), e_{j\neq i}^B(\lambda))$, it follows that if $V_i^B(\lambda) \geq U_i^F(\lambda, e_i^{DF}(\lambda), e_{j\neq i}^B(\lambda))$ (condition (1.19)) is satisfied, then $V_i^B(\lambda) \geq U_i^F(\lambda, e^B(\lambda))$ (condition (1.20)) is also satisfied. If ex-ante deviation is unprofitable, neither will be ex-post deviation; if deviation is profitable ex-post, then so will it be ex-ante.

We now summarize the results regarding a bargaining equilibrium as follows:

**Proposition 1.2.** *a bargaining equilibrium can only hold if (i) institutions are biased in favor of the player who finds itself at an advantage in $t = 1$; or if (ii) institutional bias in favor of the disadvantaged ($\overline{\lambda} + \underline{\lambda} - 1$ in the case of $L$, $1 - \overline{\lambda} - \underline{\lambda}$ in the case of $M$), difference in future advantage that is at stake ($\overline{\lambda} - \underline{\lambda}$), the importance of the future $\delta$, and the surplus shrinkage factor $\varphi$ are jointly small enough.*

## 1.4 Understanding U.S. Strike Waves through the Lens of the Model

My model explains strikes in terms of three parameters: institutional environment (bias and stakes, determined by $\overline{\lambda}$ and $\underline{\lambda}$), the importance of the future ($\delta$), and strike destructiveness ($\varphi$). I shall now discuss how the model fits historical patterns of strike activity, focusing on the U.S. experience in the last century. The time series in Figure 1.5 below, compiled by McCammon (1990) using data from the U.S. Bureau of the Census[2] and the U.S. Bureau of Labor Statistics[3], indicates there were roughly two 'waves' of industrial unrest in the 20$^{\text{th}}$ century U.S., the first spanning from the mid-1880s to 1920, the second from the early 1930s to the late 1970s. Within those waves one can identify mini-troughs roughly in the 1890s and the decade between the mid-1950s and mid-1960s. Though not shown in Figure

[2] *Historical Statistics of the United States: Colonial Times to 1970, 1975*; and *Statistical Abstract of the United States*, various years.

[3] *Analysis of Work Stoppages*, various years

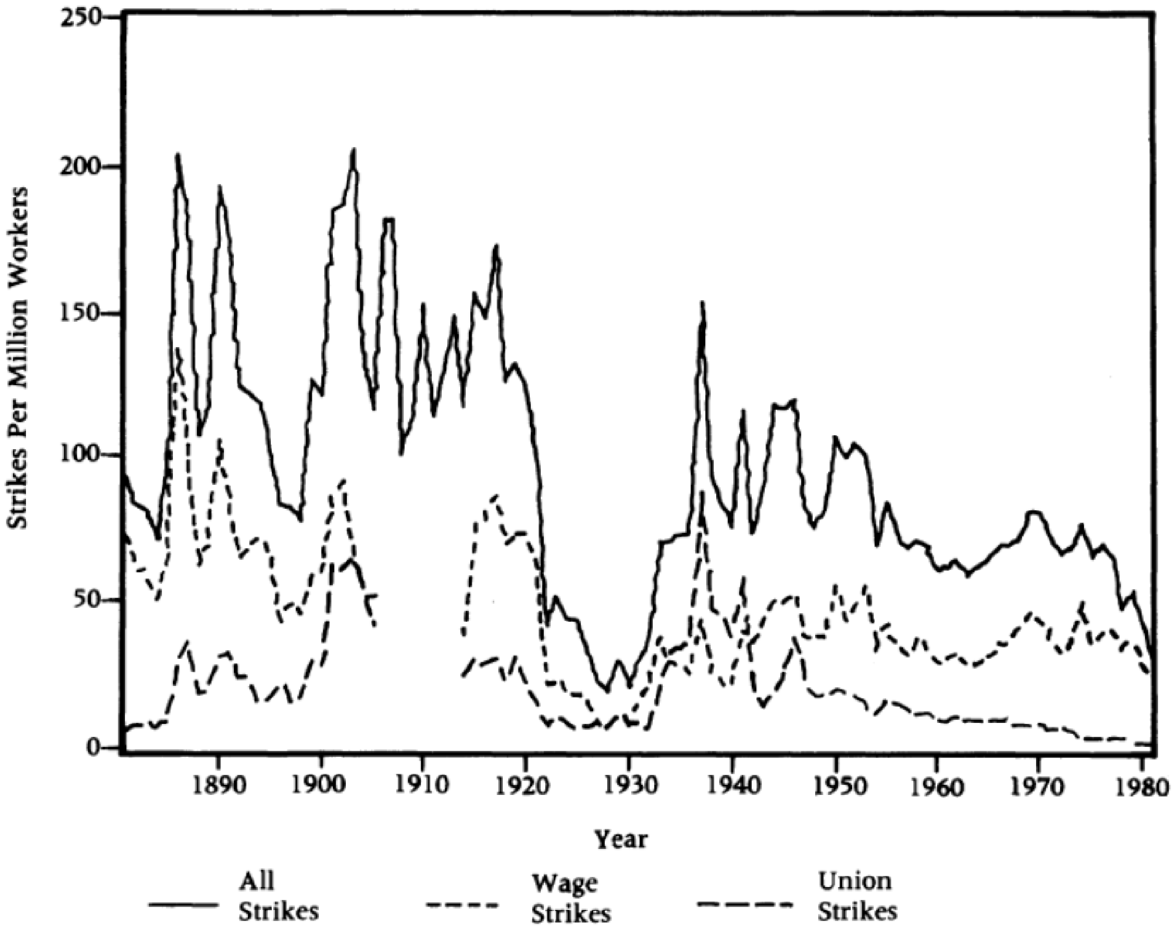



Figure 1.5: Strike Frequency, 1881-1981 (source: Fig.1 in McCammon (1990))

1.5, strike activity as measured by the number of work stoppages involving 1,000 or more workers (tracked by the U.S. Bureau of Labor Statistics) declines steeply in the early 1980s and remains very low all the way into 2017. Another visible pattern is the long-run decline in the frequency of strikes. The patterns described are not unique to the U.S.: for example, Kelly (2012) and Silver (2003) identify the 1880s as the first labor upsurge in the major western countries, followed by a slump in the 1890s, an upswing before and after World War I, a downswing in the 1920s, then an upswing in the '30s and '40s, followed by another in the '60s and '70s. Persistent low levels of strike activity since the late 1970s also appear to be a global phenomenon, as shown by Perry and Wilson (2007).

The international co-movement in strike trends strongly points to the role of economic forces in conditioning labor unrest. And indeed it would be impossible to account for the 1930s and '40s wave without reference to the Great Depression and its aftermath; nor would it be plausible to leave out the impact of globalization and the decline of manufacturing employment in explaining the post-1980 drop in strikes. In my model, the effect of economics is most directly captured by the parameter $\delta$, which can be viewed as a function of $g$, the expected growth rate of the contestable surplus, and of $\beta < 1$, the rate at which the future is discounted: $\delta = \beta(1 + g)$. The empirical measure that most accurately captures $g$ would be the growth rate of value-added by workers at the firm level, but it may also be proxied reasonably well by aggregate measures such as productivity growth or GDP growth.

Seen in this light, it is no coincidence that the first major wave of labor unrest begins around 1880: in the U.S., the Gilded Age of 1870-1900 saw a massive expansion of industry that was unprecedented in its speed and scope. Romer (1989) estimates the average annual real GNP growth rate in the period 1870-1908 at 4.61% (2.71% in per-capita terms), comparable to some of the best years of the recent few decades. It also makes sense that the second wave starts in the depth of the Great Depression and spikes in the mid-to-late 1930s: the years 1932-37 saw an average annual growth rate of a whopping 7.1%[4]. The decline beginning in the late 1970s and continuing into subsequent decades could then be attributed to the Oil Shock, the Volker Shock, and the subsequent decades of slow growth (growth rates in the 80s and 90s were not low by historical standards, but lower than the 'abnormally high' rates in the post-war boom years; see McNally, 2011). This picture is also borne out by estimated trends in productivity growth: non-farm labor productivity growth averaged 2% per year during the 1890 to 1930s period, increased to 2.34% in the 1945-72 period, and plunged to less than 1% in 1973-88 (Goldin, 1994). Thus a simplistic (and economically reductionist) account of the ebbs and flows of strikes in the past century might be that strikes surged (declined) whenever the rewards of fighting in terms of the size of the future pie to split grew (shrank).

[4] *Historical Statistics of the United States: Colonial Times to 1970*, Series F 31.

The trouble with such a simplistic account, of course, is that it has no hope of telling the whole story. For instance, it cannot explain the dip in the 1920s, even though annual growth rates in 1920-29, at 4.3%, were comparable to the preceding period[5]; neither does it explain the peace and quiet in the later part of the 1990s when productivity growth did accelerate to an annualized rate of 2.8% (which then tanked again in 2004: see Jorgenson et al., 2008, for an analysis of this 'productivity resurgence'). This is where institutional environment comes in. Although I have loosely described the space of parameters $(\overline{\lambda}, \underline{\lambda})$ as capturing institutions, a term often associated with formal rules and governmental policy, the essence of $\lambda$ is the power of labor to enforce its will upon management, either through the threat or actuality of successful collective action. Insofar as they influence the facility of such collective action, a variety of socio-economic as well as legal and political factors can be viewed as making up the institutional environment: e.g., the presence of a layer of committed rank-and-file activists that can help overcome collective action problems, political parties (or simply public opinion) sympathetic to labor's cause, civic protest movements outside the workplace, degree of social cohesion among groups of workers involved, ability of firms to relocate, and levels of unemployment to name a few. So how did these forces come into play?

Let us start from 1880. Economic growth resumed in the 1880s after a prolonged period of depression in the 1870s. Legal institutions were clearly unfavorable toward labor (biased in favor of management): although some states were moving away from the Common Law tradition of treating worker 'combinations' as criminal conspiracies in favor of recognizing the right of unions to exist, the right of employers to conduct business usually trumped workers' right to strike (Currie and Ferrie, 2000). However, other developments may have been working in favor of labor. Popular resentment of the rapidly growing inequality in the Gilded Age certainly found expression in the broad appeal enjoyed by egalitarian and socialist ideas embodied, for example, in Edward Bellamy's novel Looking Backward (Zinn, 2015). The spectacular growth in this period of the Knights of Labor, who broke from the

[5] *Historical Statistics of the United States: Colonial Times to 1970*, Series F 31.

tradition of craft unionism to encompass unskilled workers, likely reflected this zeitgeist – as did the popularity of the movement for an 8-hour working day that culminated in the May 1, 1886 general strike. The extremely high rates of workplace fatalities in the railway and other industries would also have served as an obvious rallying cause uniting workers: on the railroads some 75,000 workers perished from the Civil War to the beginning of WWI, and the construction industry itself stated that each floor of the new skyscrapers cost a worker's life (Moody, 2014).

Taken together, it is not obvious if the institutional environment of the 1880s as a whole favored either labor or management. But if we suppose it favored labor ($L$ is the top dog), my model would predict that strikes in this period were generated by unions that were initially disadvantaged – or equivalently, groups of workers who never had a formal union before and were seeking recognition as a union. This is indeed what Card and Olson (1995) conclude based on their analysis of strike outcomes in the 1880s: strikes in that period were primarily recognition strikes. Their estimate of the wage increase due to winning a strike (which forms one basis of their conclusion about the nature of strikes in that period) is much larger than estimates of the union wage differential at the peak of union membership in the U.S. (13% vs 5%), which suggests the stakes of fighting were also higher in this period.

Then came the 1890s, a decade filled with negative shocks to labor. Economically, there were the depressions of 1893 and 1896 which by themselves would have both reduced $g$ (and therefore $\delta$) and shifted institutional bias away from labor and toward management (for the simple reason that the mass of the unemployed would have supplied a readily available pool of replacement workers in the event of a strike). Legally, this was a period when the use of court injunctions to break up strikes soared. Whereas previously unions deemed to be violating employers' rights could only be prosecuted after lengthy trials, injunctions could be granted after a brief hearing and a mere assertion that harm to a firm's commerce was imminent (Currie and Ferrie, 2000). According to Naidu and Yuchtman (2016), injunctions were the "institutional innovation that most effectively pacified labor violence". Most importantly, the 1890s saw disastrous defeats of high-profile strikes such as the Homestead Strike of 1892 (broken up by the Pennsylvania state militia) and the Pullman Strike of 1894 (broken up by federal troops). The latter strike was broken with the help of an injunction; its leader, Eugene Debs, was prosecuted for violation of the Sherman Antitrust Act of 1890. In the meantime, the Knights of Labor was decimated as an organization as a result of internal fractures as well as concerted employer offensives (see Voss, 1993 for a full history of the Knights). In short, this was a period when labor was an underdog at a disadvantage.

Then something unprecedented occurred in 1902: the federal government intervened for the first time as a neutral arbiter in the anthracite coal strike in eastern Pennsylvania that year. Against vehement opposition from employers, Theodore Roosevelt imposed a settlement that granted the miners a 9-hour day and a 10% wage increase (exactly half the gains the United Mine Workers of America had demanded) which the miners regarded as a victory (Grossman, 1975). This was one episode that heralded things to come in this Progressive Era: the growth of the Socialist Party of Eugene Debs as well as that of the militant syndicalist organization Industrial Workers of the World (IWW). Nevertheless, it would be a stretch to say that institutions were generally favorable to labor in this period. Legal institutions remained hostile to unions, and employers began to effectively coordinate anti-union offensives on a national scale through the newly formed National Association of Manufacturers (NAM). NAM's 'open shop drive' that began around 1903 and lasted for a decade was by and large successful in thwarting unionization (Griffin et al., 1986). It is possible, then, that the persistent high levels of strikes in the first decade of the century was driven not so much by newly unionizing workers as by craft unions that had established themselves in the previous decades and remained in a position of advantage. The 'scientific management' movement underway that sought to break the labor process into low-skill tasks was in part an attempt by management to undermine craft unions by breaking the craftsman's monopoly of knowledge over the production process, and provoked tremendous resistance in the form of greater strike activity and high turnover (Brody, 1993; Montgomery, 1980). Thus the 1900s may have represented a fighting equilibrium in which unions were the underdog fighting to maintain their advantage.

The more significant institutional shift came after the Democratic takeover of Congress in 1910, and then again with the outbreak of World War I. The American Federation of Labor's (AFL) strategy of allying with the Democrats to win pro-labor reforms paid off in 1912, when Congress established the U.S. Department of Labor, which took a relatively progressive view of unions; in 1914 Congress passed the Clayton Anti-trust Act which limited the use of injunctions to break strikes (Wallace et al., 1988). Then came the war. The Wilson administration sought to safeguard wartime production by placating labor, in particular by favoring AFL unions in war industries – even as it violently cracked down on the IWW and the Socialist Party (Moody, 2014). In the event, swelling employment in war industries, skyrocketing inflation, and public outrage over war profiteering by industrialists led to a surge in strikes, especially among already unionized workers (McCartin, 1997). The end of the war saw massive strikes in the steel industry in 1919, a general strike in Seattle 1920 where unions took charge of the city, a virtual civil war in West Virginia and central Illinois, among other episodes. This was part of an international wave of radicalization inspired by the Russian Revolution of 1917.

The mass strikes of the early 1920s all ended in defeat for the workers, however, and the pendulum of institutional bias swung back in management's favor. With no longer any incentive to placate labor, the federal government ceased to offer protection and employers were left free to unleash a wave of anti-union campaigns. Two tactics championed by NAM that gained prominence in this period were company unions and 'welfare capitalism' (i.e. corporate welfare benefits), which were devised as 'carrots' to lure away would-be unionists (Griffin et al., 1986). In combination with the more traditional 'sticks' of yellow-dog contracts, blacklisting, etc. these tactics led to a collapse in union membership and strikes.

These developments suggest the 1920s could have corresponded to a no-strike equilibrium, where labor was both disadvantaged and institutionally marginalized. Another interpretation could be that the stakes of fighting was diminished: after all, workers' standards of living did modestly improve even in the absence of strikes (although the unskilled trades fared much worse than the skilled: see Stricker, 1983 for a breakdown of wage gains by group in the 1920s). At least some workers may not have felt the need to fight because the corporate benefits and wage increases their employers were offering already captured something akin to a 'union threat effect'. It may be argued that the specter of Bolshevism was real enough (and the recent memory of mass labor uprisings haunting enough) to bolster the baseline bargaining power ($\underline{\lambda}$) of workers (as perceived by employers), while $\overline{\lambda}$ was simultaneously lowered.

At any rate, what little wage growth the 'roaring 20s' offered workers came to an abrupt end with the onset of the Great Depression. Not surprisingly given the patterns in previous depressions, this doesn't lead at first to any appreciable uptick in strikes. Instead it was the movements of the unemployed and farmers, often led or influenced by the Communist Party, that began to effect a profound shift in the political climate (Goldfield, 1989). These broader social movements likely presaged the coming great wave of labor insurgency, but it took the beginning of economic recovery in 1933 and the passage of the National Industrial Recovery Act (NIRA) the same year to trigger a full-blown upsurge. Section 7(a) of NIRA enshrined workers' right to join a union and bargain collectively for the first time in a federal statute, which labor perceived as a signal of government support for genuine unions; employers, meanwhile, read in the same clause government approval for setting up company unions. Thus in the two years before NIRA was struck down by the Supreme Court in 1935, recognition strikes by newly emboldened workers surged in parallel with the number of company unions (Bernstein, 2010), which took on the character of a race between labor and management to establish facts on the ground in order to bend the interpretation of the law to its side. Employer violence only fueled more worker radicalism. The breakthrough strikes of 1934 in Toledo, Minneapolis, and San Francisco, all led from the grassroots by socialists of various stripes, impressed upon the political establishment that the alternative to reform could well be revolution (Goldfield, 1989).[6] This was likely a decisive impetus behind the passage of the National Labor Relations Act (NLRA, or Wagner Act) in 1935, and more broadly the Roosevelt administration's 'Second New Deal' (Selfa, 2012). The even more militant sit-down strikes at General Motors and elsewhere in 1936-7 (again led by Communists) probably forced the Supreme Court to affirm the constitutionality of NLRA in 1937. From this point, employers retreat from open resistance and turn to lobbying to amend the NLRA (McCammon, 1990).

Whether the NRLA (which forms the legal foundations of present-day labor relations) actually benefited or impeded the labor movement has long been a subject of debate, particularly between industrial pluralist scholars and critical legal scholars. Following McCammon (1990), I argue that it served to simultaneously strengthen and weaken organized labor, by legally authorizing collective action and thereby subjecting them to regulation by the state. In terms of my model, the function of the NLRA was to increase $\underline{\lambda}$ by offering workers a measure of protection from the worst anti-union practices of employers, while also lowering $\overline{\lambda}$ by subjecting strikes to more rigid formal procedures – thereby making them both more predictable to management and amenable to control by the more moderate union officialdom. This should in theory have had the effect of promoting peace by reducing the stakes of fighting, and it appears the intent of the law, at least, was to reduce the high levels of strikes (Klare, 1977). As stated in Section 1 of NLRA: "Experience has proved that protection by law of the right of employees to organize and bargain collectively safeguards commerce from injury, impairment, or interruption..." In the short run the legislation had, if anything, the opposite of the intended effect: a working class revolt that escaped all bureaucratic control.

[6] As one House representative speaking in support of NLRA in 1934 warned: "You have seen strikes in Toledo, you have seen Minneapolis, you have seen San Francisco... but you have not yet seen the gates of hell opened, and that is what is going to happen from now on [if the NRLA isn't passed]." (quoted in Goldfield, 1989).

It may have been that in the specific context of the time, the mere suggestion that the state was willing to come to the aid of labor (as first signalled by NIRA) was enough to embolden workers to the point it became impossible to cap their power below a desired $\overline{\lambda}$. It may also be that the extreme intransigence of employers, more than the New Deal policies themselves, provoked workers' radicalization. At any rate, the broad picture of the 1930s corresponds to a fighting equilibrium in which advantage shifts on a massive scale from management to labor, which is simultaneously enabled by a shift in the institutional environment and reacts back to accelerate that shift in the institutional environment. The advent of World War II put a temporary break on labor militancy as the Congress of Industrial Organizations (CIO) sought to restrain the demands of its members to aid the war effort. The war's end brought explosions of pent-up grievances in the form of wage strikes involving millions and general strikes in six different cities (Moody, 2014).

The first serious reaction of the employer class to this unpleasant state of affairs came in the form of the Labor Management Relations Act (the Taft-Hartley Act) of 1947. This piece of legislation amended the Wagner Act in a way that severely limited the right to strike: among other things, it made sympathy strikes and wildcat (mid-contract) strikes illegal (whenever a contract contains a 'no strike' clause); denied the right to vote in union elections to striking workers for whom permanent replacements have been hired; and limited the types of issues over which a union can legally strike to 'mandatory issues' pertaining to wages, hours, and other terms of employment – effectively putting a wide range of workplace control-related issues out of the negotiating table (McCammon, 1990). An immediate consequence was the advent, for the first time, of union de-certification elections in 1948, instigated by employers wanting to take advantage of the fact that strikebreakers could (but strikers could not) vote to de-certify an existing union: of the 97 de-certification proceedings conducted by the National Labor Relations Board that year, 63 were successful (NLRB Annual Reports 1936-49). Restricting the scope of bargaining issues and outlawing wildcat and sympathy strikes was directly aimed at undermining the very feature of strikes up to that period that gave them such enormous appeal to workers: the power it afforded workers to exercise control in the workplace, and the solidarity it attracted from other groups of workers. A steady stream of court cases and new legislation in the 1950s and '60s reinforced such limits on the form and content of the strike weapon (McCammon, 1990). Alongside this sustained downward pressure on $\overline{\lambda}$, the purging of Communists in the McCarthy years of the 1950s deprived the labor movement of the layer of activists that had played such a pivotal role in the upsurge of the '30s and '40s. These favorable institutional developments (from the employers' perspective) and the recession in 1957 explains the 'Management Offensive of 1958-63' (Davis, 1999), in which employers such as General Electric and U.S. Steel felt confident enough to bargain aggressively to remove the power of workplace representatives in their contracts. But the power of organized labor was still too entrenched to shake off with a single offensive. As employers embarked on a drive to enhance productivity through work intensification in the 1960s, strikes surged again (many of them wildcat) until they peaked at over 6,000 strikes in 1974 (Moody, 2014).

The period from the late 1960s to the late 70s could be characterized as a fighting equilibrium where unions were the underdog fighting to hold on to its advantage. But this advantage could not be sustained indefinitely. The two recessions of 1973-75 and 1980-82, along with devastating defeats such as those of the coal miners in 1978 and the PATCO air traffic controllers in 1981, inflicted crippling damage to labor. The decade that followed saw a resurgence in the use of permanent replacement workers and labor injunctions, anti-strike weapons that had been legally available but infrequently used in previous decades (Moody, 2014).[7] Private-sector union membership, NLRB election win rates, and strike frequency all declined precipitously. U.S. labor entered a period of no-strike equilibrium similar to the 1920s. But this era of industrial peace proved much longer-lasting than the 1920s: after almost three decades of near-continual decline, work stoppages in the U.S. has yet to recover

[7]Labor injunctions, which had been prohibited by the Norris-LaGuardia Act of 1932, was revived by the Supreme Court in the 1970 *Boys Market* case.

from their historically low levels.

A number of factors might plausibly account for this exceptional duration. First, it must be noted that in a dynamic sense, no-strike (bargaining) equilibria are inherently more stable than striking (fighting) equilibria: the former constitute the absorbing state into which unions and firms starting out in the latter state will eventually enter, as soon as an underdog loses a battle and thereby becomes disadvantaged. What perturbs a no-strike equilibrium is either a shift in the institutional environment that turns underdogs into top dogs, or an increase in the fighting capacity of the underdog that may lift it into a position of advantage. In the 1930s it was arguably the National Recovery Act and the subsequent series of New Deal legislation that shifted the institutional terrain and thereby opened the floodgates of industrial unrest. Such a Polanyian counter-movement toward social protection did not arise in the post-Reagan period. Even the Obama administration's response to the Great Recession did not include any serious support for organized labor (Lichtenstein, 2013). It is also possible to speculate that a confluence of factors such as globalization, shrinking share of manufacturing employment, and dominance of free-market ideology since the demise of the USSR has conspired to entrench labor's disadvantaged position.

The year 2018 offered the first signs of a possible perturbation in the prolonged no-strike equilibrium: the number of workers participating in work stoppages that year hit the highest on record since 1986, driven by public school teacher strikes and walkouts spreading from West Virginia to Oklahoma, Arizona, and California. Many of the key organizers behind those teacher strikes had been radicalized and jolted into union activism by Bernie Sanders' insurgent 2016 campaign for the Democratic Party's presidential nomination (Blanc, 2019). Such an influx of ideologically motivated activists into hitherto disadvantaged unions strengthens their fighting capacity and may thereby induce a fighting equilibrium where the underdog has the advantage. If 2018 proves to be a turning point in long-term strike trends in the U.S., part of the explanation would lie with the transformed terrain of political and economic discourse effected by Bernie Sanders' message of 'democratic socialism'.

In my discussion of historical patterns of strikes, I have thus far left out the role of the parameter $\varphi$. Strike destructiveness can be thought of as a function of contract length and strike duration. Since both of these involve choices made by the contestants, $\varphi$ is not exogenous in the way I treat it in the model. I leave it to future work to endogenize $\varphi$ (or strike duration) in the setting of a rational strike with complete information.

Also left in the background is the role of union officials in promoting or suppressing rank-and-file activity. Although the distinction between the union officialdom and the rank and file is an important one, and tensions between the two a recurring theme in labor history, I treated them as a homogeneous entity called 'labor' or 'union'. To the extent union leaders undermine the militancy and self-activity of the rank and file, they can be viewed as biasing the institutional environment against labor; to the extent they promote the efficacy of collective action, they can be viewed as doing the opposite. The complex interaction between the union leadership, the rank and file, and management is a subject to be tackled in the next chapter.

## 1.5 Conclusion

I have demonstrated that a strike can be a rational strategy for both labor and management when the stakes of fighting, in terms of the gain in the winner's future advantage, is high enough, and if the institutional environment is biased in favor of the party that stands to increase future advantage by winning today. The advantage of my model relative to existing models of rational strikes is that it does not rely on assumptions of information asymmetry, and allows agents to consider the impact of the outcome of a strike or bargaining over the current labor contract on future contract bargaining. Needless to say, a theory of rational strikes under complete information does not preclude other explanations: information asymmetries, 'irrational' fear or feelings of aggrievement, faulty bargaining, etc. all likely play a role in the making of actual strikes. But there is a sense in which grounding strikes on informed rationality renders all other explanations as rather secondary: it reveals that conflict is something deeply embedded in the employment relationship and not easily removed by refining the bargaining institution.

My model fits some of the broad patterns observed in the history of U.S. strike waves: the general pro-cyclicality of strikes; the onset of strike waves coinciding with changes in the institutional (sociopolitical) environment that favor the previously disadvantaged underdog; and conversely, the decline of strike waves following adverse institutional changes for the disadvantaged side. A limitation of the model lies in its simplifying assumption that strikes are a winner-take-all contest that can only have two distinct outcomes, win or loss, which in turn set the union's advantage parameter to one of two exogenously given levels. This feature prevents analysis of intermediate cases such as strikes ending in compromise or the case of unions with initial advantage lying somewhere between the two given values. Moreover, the model itself says nothing about how specific elements of the broadly defined 'institutional environment' (unemployment rate, labor laws, level of inequality, dominant ideological discourse, prevalence and success rate of strikes, etc.) affect the fighting capacity of individual unions, although I assume they affect the latter in particular ways in analyzing the history of U.S. strike waves. The task of theoretically articulating the relation between such environmental factors and collective action capacity is left for future work.

# Chapter 2

# Iron Law of Oligarchy? Understanding the Behavior of Union Officials

## 2.1 Introduction

The basic mission of labor unions is to fight for improvements in pay and conditions for workers. A huge literature has demonstrated that unions indeed deliver higher pay and other benefits for workers they cover (see, for example, Freeman and Medoff, 1984 and Blanchflower and Bryson, 2004). Revitalizing labor unions is also increasingly offered as a policy prescription to combat rising income and wealth inequality (e.g. Reich, 2016), which may have some merit in light of evidence that union decline in the U.S. in the decades since the mid-1970s has non-trivially contributed to rising inequality (Farber et al., 2018; Western and Rosenfeld, 2011; Card, 2001; DiNardo and Lemieux, 1997).

At the same time, however, unions – or more precisely, their leaders – often behave in ways that seemingly run counter to the interests of the workers they claim to represent, thereby discrediting themselves in the eyes of would-be unionists and the public at large. Sometimes they appear to capitulate to the demands of employers without a fight: an example in the U.S. might be the early-1980's turn toward 'concession bargaining', whereby national unions agreed to wage cuts and/or other onerous terms under threat of business restructuring, despite sometimes bitter objections from local-level leaders and rank-and-filers (Craft et al., 1985; Slaughter, 1983; Moody, 1988). In periods of heated labor conflicts, union leaders were frequently seen as siding with management to end or sabotage strikes, even when little of the workers' demands had been won (Brecher, 2014). The internal regimes of unions have been frequently authoritarian (in substance if not in form), which allowed many leaders to crush dissent and to hold nearly lifetime tenures as heads of their organizations (Friedman, 2007). In extreme cases, union officials have acted as willing accomplices in labor racketeering schemes where organized criminals infiltrated unions to embezzle union funds, enforce employer cartels, and secure favorable terms for employers in labor contracts in exchange for bribes – using violence and intimidation to suppress internal dissent (Jacobs, 2007).

That union officials elected to serve their members regularly behave in this way is a phenomenon that cries out for an explanation. Simple attributions to human corruptibility are unsatisfying given that union leaders are often selected on the basis of their perceived moral courage and personal sacrifice. This chapter will argue that the tendency of union leaders to disappoint their members stem from a fundamental difference in the goals pursued by the union bureaucracy and the rank and file. Using a formal game-theoretic approach and drawing from the insights of Michels (1915) and other labor movement intellectuals, I show that when union leaders maximize membership rather than some measure of worker welfare, they are prone to accept a poorer compromise for the workers than is justified by the capacity of their union to extract concessions through a strike, actual or threatened. Intuitively, this is because a confrontation with the employer always carries the risk of a defeat serious enough to uproot the union, and in the minds of union officials this danger looms larger than the potential gains to rank-and-file workers that confrontation can bring. Rank-and-file workers may feel underserved by their timid leadership, but lacking a viable alternative, enough workers may still choose to remain with the union, which affords leaders a measure of relative impunity for their actions.

The remainder of this chapter is organized as follows: section 2.2 reviews previous attempts to theorize union bureaucracy, section 2.3 presents my model of union bureaucracy, and section 2.4 concludes.

## 2.2 Institutional and Theoretical Background

Labor unions for the most part are formally democratic organizations. Not only are their leaders usually elected and accountable to conferences of delegates, members can always vote with their feet by leaving if they disapprove of the conduct of the union (unless her workplace is closed-shop, an institutional setting that ceased to exist in the U.S. since the 1947 Taft-Hartley Act). It is therefore a natural starting point to view the union leadership as simply representing the interest of its members. This has been much the standard approach in economic theory, where the union leadership was assumed to maximize the objectives of identical members (Dunlop, 1944; McDonald and Solow, 1981; Oswald, 1992), or of the median voter in the presence of heterogeneity in member preferences (Blair and Crawford, 1984; Booth, 1984; Carruth et al., 1986).

However, labor movement scholars and practitioners have long recognized that the union leadership, especially its bureaucracy consisting of full-time salaried officials, is a distinct layer inside the the union whose aims quite often came in conflict with those of the rank and file. Thus the renowned Fabian socialists Sidney and Beatrice Webb related the following

19$^{th}$-century account of the transformation of a lay union member into what they termed a 'salaried official' of the union (Webb and Webb, 1920):

> And now begins a change which may possibly wreck his whole Trade Union career [once promoted] to a salaried office... Whilst the points at issue no longer affect his own earnings or conditions of employment, any disputes between his members and their employers increase his work and add to his worry. The former vivid sense of the privations and subjection of the artisan's life gradually fades from his mind; and he begins more and more to regard all complaints as perverse and unreasonable.
>
> With this intellectual change may come a more invidious transformation. Nowadays the salaried officer of a great Union is courted and flattered by the middle class [meaning the propertied class in the context of the time]. He is asked to dine with them, and will admire their well-appointed houses, their fine carpets, the ease and luxury of their lives... With the habits of his new neighbors he insensibly adopts more and more of their ideas...
>
> Gradually he finds himself at issue with his members... A great strike threatens to to involve the Society in desperate war. Unconsciously biased by distaste for the hard and unthankful work which a strike entails, he finds himself in small sympathy with the men's demands, and eventually arranges a compromise on terms distasteful to a large section of his members... At his next appearance before a general meeting cries of "treachery" and "bribery" are raised. Alas! it is not bribery. Not his morality but his intellect is corrupted.

The Webbs welcomed the emergence of the union bureaucracy since the late 19$^{th}$ century for their conservative, moderating influence. But this same conservatism (or pro-employer bias) has been the object of much resentment among rank-and-file workers over the past century, sometimes expressed in open revolt inside established trade unions.

The perceived failure of existing trade unions and their bureaucratic officialdom to adequately represent worker interests served as a major impetus behind the growth of militant syndicalist movements in many countries in the early 20$^{th}$ century, especially in Britain where the wartime Shop Stewards' Movement mobilized the rank and file in direct opposition to their union officials' policy of cooperation with war production efforts (Darlington, 2008).

In the U.S. during WWII and the post-war decade, the leaders of the Congress of Industrial Organizations (CIO) worked hard to undermine the system of shop-floor representatives that protected workers against the tyranny of lower-level managers at the workplace, a system of bottom-up worker empowerment that had played a central role in the CIO's own success (Moody, 2010; Lichtenstein, 2013). From the late 1960s to the late 1970s, the U.S. in particular saw an explosion of rank and file-initiated 'wildcat' strikes – unofficial strikes taking place during the term of a labor contract that usually includes a no-strike clause – that were met with the determined resistance of not just employers but also union leaders, who sometimes hired goons wielding baseball bats to break up unauthorized pickets (Winslow, 2010b). More than a third of all strikes in the U.S. between the late 1960s and early 70s were unofficial, and probably a good number of official strikes were forced upon an unwilling leadership by the rank and file (Winslow, 2010b). The issues at stake, among others, were wages that barely kept up with inflation, speedup of work, erosion of work rules, and deterioration of workplace health and safety – all results of concessions made by union officials in the face of a management offensive starting in the late 1950s to ramp up productivity (Davis, 1999; Brenner, 2010).

In the more recent past, open rank-and-file revolts against union leaders have been rare, in part because the end of the Long Boom and the series of economic crises that followed have sapped the confidence of lay workers to challenge either their employers or their union bureaucracies. But tensions between the rank and file and the leadership did occasionally resurface at flashpoints of labor militancy, such as when UAW members at Chrysler voted down an agreement pushed by the leadership in 2015 to preserve the hated two-tiered wage system that the UAW had accepted during the Great Recession[1]; and in 2017 when the striking West Virginia teachers defied their union president's call to return to work before their demands had been won (Blanc, 2019).

A number of attempts have been made to theoretically explain the apparent tendency of union leaders to disappoint their followers. Michels (1915) famously claimed that all organizations, even working class parties pursuing the most democratic ideals, inevitably tend toward oligarchy – that is, domination by a minority of leaders who escape all accountability (the so-called 'Iron Law of Oligarchy'). Michels was the first to spell out the material conditions that can undermine democratic control of the masses over their leaders, namely the monopoly of specialized knowledge and skill that the leaders come to acquire as organizations grow in size and complexity. However, erosion of accountability is not enough, by itself, to explain a systematic tendency toward betrayal. The leadership must have its own set of interests counterposed to those of the masses in order to betray the latter, aided by their unchecked power. For Michels, it was the preservation of the organization (which forms the material basis of the leaders' oligarchal privileges) that became the overriding concern of the leadership. Leaders thus acquired a tendency to shy away from confrontations with the employers or the state, which might advance the interests of workers but at the cost of jeopardizing the survival of the organization. This tendency was also noted by Michels' Marxist contemporaries such as Rosa Luxemburg and Leon Trotsky (Eidlin, 2019).

Michels (1915)'s analysis has been hugely influential and informs much sociological work on union democracy (see Voss, 2010 for a summary of empirical work that sought to prove or disprove the Iron Law of Oligarchy as applied to unions). But his view is one-sided in

[1]Bradbury, Alexandra (2015, October 1). Chrysler Worker Vote 2 to 1 to Reject Two-Tier Pact. *Labor Notes*. Retrieved from https://labornotes.org/2015/09/so-far-chrysler-workers-roundly-rejectingtwo-tier-pact.

that it downplays the possibility of powerful union leaders actually leading a serious fight; it ignores the reality that strikes are often called from the top down and won with the aid of organization and leadership. Moreover, it is not entirely clear that abstention from battle is the best way to ensure organizational survival. Faced with an employer offensive to take back hard-won worker gains, for example, cowardice before the enemy could just as easily lead to mass desertion of membership and to eventual demise of the union.

Writing in the late 1940s when organized labor still seemed a radical force, C. Wright Mills (1948) in his classic sociological study of U.S. union leaders offered a more optimistic interpretation of their oligarchic face. For Mills, union leaders functioned as parliamentarians and army generals at the same time. Their authoritarian character was born of necessity to enforce discipline among the ranks of workers when engaging in battle against management; discipline also had to be brought to bear against discontented elements in their ranks to uphold whatever agreement the leaders might reach with the employer. Lacking this power to police internal discontent, union leaders would lose credibility as responsible negotiating partners; to safeguard the stability of the collective bargaining institution that ultimately serves their constituency, it becomes necessary for union leaders to act as 'managers of discontent', a shock absorber for both the rank and file and the company management (Mills, 1948).

This view presumes that union leaders do serve the interest of workers, albeit in a more or less heavy-handed fashion. It does not envision the possibility of leaders putting the stability of the collective bargaining institution (or of the union organization) above the welfare of the rank and file, as Michels and other labor intellectuals had observed.

Hyman (1975) and Cliff and Gluckstein (1986), from a Marxist perspective, offered a more 6nuanced analysis by focusing on the role of the union bureaucracy as mediator between conflicting class interests. Full-time union officials occupy a unique position in capitalist society, being neither dependent on wages paid by the employers nor on the labor of others.

Instead they derive their status and income from being seen as indispensable middle men whose function is to negotiate the terms of employment for the mass of workers. They have to be deemed indispensable to the workers in that they deliver some improvements to their livelihoods; but they also have to appear indispensable to employers in that they are capable of dampening workers' demands and suppressing worker militancy more effectively than employers themselves could – which is often true because of union leaders' credentials as former militants. The union bureaucracy is thus caught between a rock and a hard place: if they side too closely with the employers, they lose authority over their members (and possibly lose membership, or else get voted out of office); but if they push the demands of the membership too hard, their raison d'être as 'responsible' mediators will be called into question by employers and the state. A violent confrontation may ensue, in which the bureaucracy risks having its assets confiscated and being imprisoned, while losing all control over the rank-and-file militants who now take matters into their own hands. The balance of these two competing pressures can push the leadership into one direction or the other at different times, but the leadership will always be careful not to overstep the bounds of what is acceptable to employers and the state unless they absolutely have to.

What is lacking in the above account is a precise formulation of the union bureaucracy's objective. They may seek 'status' as mediators, but how is this measured and how do factors such as the balance of forces between labor and capital affect this objective?

A handful of economists have provided a clue to answering this question, namely that union officials seek to maximize the size of their organizations (Ross, 1956; Atherton, 1973; Martin, 1980; Farber, 1986). This is an eminently reasonable conjecture, since the size of a union not only affords leaders prestige and renumeration but also a reason to be taken seriously by employers as negotiating partners. Moreover, insofar as the collective power of workers to extract concessions from employers depends on the size of unions, union officials who prioritize membership building can claim to serve the interest of the workers without actually exercising that collective power to the fullest extent.

The trouble is that the determinants of union size (i.e. membership) is not very well understood. Dunlop (1944) first suggested an arbitrary membership function that is increasing in the wage set by the union. Lewis (1963)'s model of a 'boss-dominated' union posited that workers chose to join if the expected utility of getting a scarce union job exceeds the utility of a non-union job; Farber (1986) further developed the idea by introducing heterogeneity in workers' non-union wages. These models of membership determination assume a setting where either the union controls labor supply (as in industries dominated by hiring halls) or all employed workers become members (as in a closed-shop firm), and are not generalizable to modern settings. The problem of recruiting new members among those who are already employed (and covered by a union contract) is ignored.

With these considerations in mind, I proceed in the next section to develop a formal model of union behavior built on the premise that union leaders seek to maximize union membership, which also specifies how workers choose to become members when they have the option to free-ride. Treating union membership as akin to an act of voluntary contribution to a public good, the model generates a membership function that is increasing in the benefit the union brings to workers. From this it may appear as though the goal of the union bureaucracy aligns with that of the workers, but it will be shown that under certain conditions their goal will conflict with what is best for workers. It is hoped that the formal model will offer a rigorous explanation for the leadership vs. rank-and-file conflict observed throughout labor history; however, for the sake of mathematical tractability some of the richness of sociological insight present in the works of the labor movement intellectuals cited in this section will inevitably be forfeited.

## 2.3 A Model of Union Bureaucracy

This section consists of three parts. Subsection 2.3.1 develops a baseline model of collective bargaining and strikes that takes into account the three-way interaction among the union leadership (union bureaucracy), the employer, and rank-and-file workers. It will be seen that the union leadership's membership-maximizing objective can come in conflict with the interests of the rank and file especially when the collective action capacity of the union is strong. Subsection 2.3.2 considers an extension in which the employer can co-opt or effectively bribe the union leadership by offering enhanced union security in return for moderating the workers' demands. Subsection 2.3.3 extends the baseline model by letting union power depend on membership, and considers the resulting dynamic co-evolution of union membership and union power.

### 2.3.1 The baseline model

#### Preliminaries

We assume an institutional setting where union membership is voluntary (i.e. no 'closed shop' agreements), and where a single union exclusively represents all workers in a unionized firm (this is the established system in the U.S.). There are three agents in the model: worker $i \in [0,1]$ in a firm employing a unitary mass of workers; management $M$; and union leadership (bureaucracy) $L$. $M$ and $L$ bargain over how to split a contestable economic surplus of size $s$. $L$ can secure a part of the surplus, $\alpha < s$, either through a peaceful settlement or through a strike. $\alpha$ can be interpreted as the union wage premium; hereafter I will call it the workers' 'share of the surplus', although strictly speaking it is not a ratio but an absolute magnitude. Upon observing the outcome $\alpha$, workers individually decide whether to join the union.

The objective of $L$ (union leadership) is to maximize union membership, $m$. But this requires workers to voluntarily sign up to become union members despite the positive marginal cost associated with it. This cost $c < s$ (which we assume to be fixed for now) includes not just membership dues but also the possibility of being discriminated against in promotions, having to participate in union events under threat of shame, etc. If workers have preferences that are purely self-regarding and if the marginal contribution of each member to the collective cause is vanishingly small (as would be the case when individual workers are a point mass in an interval), then every worker would choose to free-ride and $m$ would equal zero. This is clearly not the case in the real world, and so we make a non-standard assumption about worker preferences in order to derive a union membership function, namely that workers obtain varying degrees of satisfaction from belonging to an entity that benefits his/her community of co-workers. We thus specify worker utility in the following way:

$$U_i = (i \cdot \alpha - c)\mathbb{1}\{join\} + \alpha, \forall i \in [0, 1]$$

Note that $i$ in this specification serves a dual purpose: it captures the heterogeneous degree that workers 'identify' with or derive utility from belonging to an organization that brings about a positive social benefit (in proportion to the magnitude thereof) as well as indexing individual workers. A worker with $i$ close to 1 can be thought of as the more pro-social type in the workplace. To the extent that signing up for union membership can be viewed as an act of voluntary public good provision, the $i \cdot \alpha$ term plays a similar function as the 'warm glow' effect of giving or other pro-social preferences posited in the literature on private provision of public goods (Margolis, 1984; Sugden, 1984; Andreoni, 1990); it is also similar to the component of a voter's utility in Hinich (1981)'s model of voting, in which voters derive gratification from her act of voting for a party she believes will win.

A worker will only choose to join if $i \cdot \alpha - c \geq 0$. This means that any worker indexed by $i \geq i^*$ such that $i^* = c/\alpha$ will join the union. This immediately yields the membership function:

$$m = 1 - i^* = 1 - \frac{c}{\alpha} \geq 0$$

Note that $m$ ('membership', or 'member share') is the outcome of an already realized $\alpha$ resulting from a strike or settlement. Thus $L$'s problem is to negotiate or impose a contract ($\alpha$) with a view to maximizing membership over the term of the contract, with the knowledge that workers will reward a favorable contract with a high sign-up rate. The management ($M$) wishes to maximize the following objective:

$$U_M = s - \alpha - b \cdot \mathbb{1}\{union\},$$

where the indicator function evaluates to 1 if the firm is unionized. Thus as long as a union exists, management incurs a fixed cost equal to $b$, which can be interpreted as the cost of managing a business with less than full dictatorial control. No matter how purportedly business-friendly, a union always wants a say in the day-to-day running of the workplace and can be driven to challenge managerial decision-making insofar as it affects the well-being of employees. It need not be the leadership's intention to undermine managerial control: the very existence of a union can embolden rank-and-file workers to challenge the authority of shopfloor managers despite the leadership's commitment to safeguarding management prerogatives (see Montgomery, 1980 for a classic study of workers' attempts to defend union work rules against both their leaders and their employers). $M$'s union distaste parameter $b$ may vary from employer to employer due to technical or market conditions that differentially reward managerial authority: $b$ could be high, for example, in industries characterized by rapid product cycles where swift decision-making and implementation is at a premium. Alternatively, a higher propensity for rank-and-file militants to contest workplace control could increase $b$.

If a strike occurs, we assume for simplicity that only two outcomes are possible: either $L$ demands the entire contestable rent (i.e. $\alpha = s$) and wins with probability $\lambda$ (assumed to be public information), or loses and gets nothing. I will refer to $\lambda$ as the 'fighting capacity' or 'collective action capacity' of a union: it captures the ability of the union to hold together a strike until it forces management's capitulation.

In practice $\lambda$ will depend on membership to some degree, but for the moment I will treat it as exogenous to $m$, because it is quite possible for a union with a broad membership to have lost its ability to deploy the strike weapon through many years of non-use; it is also possible for a union with a small membership base to punch above its weight by virtue of its militancy or technological conditions that allow a small minority to bring production to a halt. I will later extend the model by incorporating initial membership as a determinant of $\lambda$.

**Equilibrium under a membership-maximizing leadership**

To determine the equilibrium outcome of this game, we begin by considering the possible consequences of a strike.

If a strike loses with probability $1 - \lambda$, all members desert and the union ceases to exist ($m = 0$). Therefore $M$'s expected utility from provoking or taking a strike is:

$$EU_M|_{strike} = (1 - \lambda)s - \lambda b.$$

In the event of a peaceful settlement, $M$ gives up $\alpha$ of the rent $s$ and the union continues to exist. Therefore $M$'s utility is:

$$U_M|_{settle} = s - \alpha - b.$$

Therefore $M$ can credibly threaten to take a strike if $\alpha$ demanded by $L$ is such that $(1-\lambda)s - \lambda b > s - \alpha - b$, i.e. if $\alpha > \lambda s - (1-\lambda)b$. This means that $M$'s reservation offer, i.e. the largest $\alpha$ that $L$ can hope to win without recourse to a strike, is given by:

$$\overline{\alpha} = \lambda s - (1-\lambda)b > 0. \tag{2.1}$$

The positivity constraint for $\overline{\alpha}$ implies that $\lambda > \frac{b}{s+b}$: otherwise, $L$ has no hope of extracting any concession at the bargaining table and would be forced to strike. $M$ would rather incur a strike than concede anything to a union whose fighting capacity is sufficiently weak, so the union might as well fight and die trying rather than face certain death. Whether such a union can pull off a strike in practice is a different matter.

Next, we proceed to derive $L$'s reservation offer $\underline{\alpha}$, or the lowest $\alpha$ that $M$ can hope to impose without provoking a strike. $L$ would prefer to strike than to accept any offer $\alpha$ that yields a smaller membership level than a strike can bring about. Thus $L$ strikes if $\lambda(1 - c/s) > 1 - c/\alpha$. $L$'s reservation offer is then:

$$\underline{\alpha} = \frac{c}{1 - \lambda(1 - c/s)}. \tag{2.2}$$

It can be checked that $\underline{\alpha}$ is between zero and $s$ for all values of $\lambda \in [0, 1]$.

As long as $\underline{\alpha} \leq \overline{\alpha}$, the two sides have room for reaching a negotiated settlement at $\alpha = \hat{\alpha}$, which would be some convex combination of $\underline{\alpha}$ and $\overline{\alpha}$:

$$\hat{\alpha} = \beta\overline{\alpha} + (1-\beta)\underline{\alpha}, \quad \beta \in [0, 1]. \tag{2.3}$$

A strike is unavoidable, however, if $\underline{\alpha} > \overline{\alpha}$, which amounts to the following condition:

$$\frac{c}{1 - \lambda(1 - c/s)} > \lambda s - (1-\lambda)b. \tag{2.4}$$

After multiplying both sides by $1 - \lambda(1 - c/s)$ and solving for the roots of $\lambda$ that turn (2.4) into an equality, it can be shown that (2.4) is satisfied if:

$$\lambda < \frac{(b+c)s}{(s+b)(s-c)} \equiv \hat{\lambda}. \tag{2.5}$$

A curious implication of this result is that it would be the weaker unions that will be tempted to strike (and whose employers would be tempted to take a strike), not the stronger ones with fighting capacity $\lambda$ above the threshold $\hat{\lambda}$ in the right hand side of (2.5).

**Comparison of worker welfare under membership-maximizing vs benevolent union leaderships**

Does the ability of stronger unions (with $\lambda > \hat{\lambda}$) to settle peacefully work to the benefit of workers? To see whether it does, it would be useful to first establish a benchmark regarding what $L$ would do if it were a 'benevolent' leadership, trying to maximize either the total wage bill or the utility of the workers as in standard models of union behavior. A wage bill-maximizing $L$ would prefer to strike whenever the expected share of the surplus to be gained from a strike, $s\lambda$, exceeds any share that can be obtained through settlement. In my model the upper limit of the share obtainable through settlement is given by $M$'s reservation offer, $\lambda s - (1-\lambda)b$, which is always less than $\lambda s$. Thus a wage bill-maximizing $L$ would always prefer to strike.

For $\lambda < \hat{\lambda}$, then, $L$ is compelled to strike in either the benchmark case or membership-maximization case, so that workers can expect the same wage differential $\alpha = \lambda s$. For $\lambda \geq \hat{\lambda}$, however, a wage bill-maximizing $L$ chooses to strike while a membership-maximizing $L$ chooses to settle, the former delivering an expected share of $\lambda s$ and the latter delivering $\hat{\alpha}$ for the workers. Thus the relative benefit to workers, in terms of wage gain, from having a membership-maximizing leadership is:

$$\hat{\alpha} - s\lambda = \beta \underbrace{[\lambda s - (1-\lambda)b]}_{=\overline{\alpha}} + (1-\beta) \underbrace{\left[\frac{c}{1-\lambda(1-c/s)}\right]}_{=\underline{\alpha}} - s\lambda. \tag{2.6}$$

This ‘benefit’ is negative even under the best of circumstances for the union. Recall that for $\lambda \geq \hat{\lambda}$, $M$’s reservation offer $\overline{\alpha}$ is higher than the union’s reservation offer $\underline{\alpha}$. Then the benefit in (2.6) is maximized when $\beta = 1$, i.e. when $L$ has all the bargaining power. But even at its maximum, the benefit equals $-(1-\lambda)b < 0$. In other words, once a membership-maximizing union bureaucracy becomes powerful enough to extract concessions without a strike, it begins to sell workers short by at least $(1-\lambda)b$, compared to the wage bill-maximizing benchmark.

Here we can see a tension developing between the interests of the union bureaucracy and the rank and file. This tension stems from the fact that for $L$ it is the preservation and growth of the union apparatus (sustained by membership) that is the end in itself, while for the workers the union is a means to an end, namely of maximizing gains for employees. These two goals begin to diverge when $L$ has to compensate $M$ for its aversion to power-sharing in order to avoid a confrontation that may put the existence of the union at risk. Somewhat counterintuitively, such tension only arises for unions that are powerful enough. This result is consistent with the historically observed pattern of once-aggressive unions growing increasingly moderate the more they become established (e.g. the AFL and CIO in the U.S.; the CGT of France).

That a wage-bill maximizing $L$ would always prefer to strike is a bit unrealistic, and is an artifact of ignoring the costs associated with strikes, such as lost income over the duration of a strike, anxiety over the possibility of getting permanently replaced, etc. Let us now introduce this cost into the analysis.

Let $\xi < s - c$ denote the cost of strike borne by workers. We assume this cost is common to all workers (regardless of union membership) and detracts from any concession $\alpha$ won by $L$. For simplicity, we ignore strike costs borne by $M$, as this adds little insight to the model. After a victorious strike, workers value the pie by $s - \xi$, and make union membership decisions accordingly; $M$'s problem remains the same as before. Thus a membership-maximizing union leadership, denoted $L^m$ hereafter, chooses strike whenever:

$$\lambda\left(1 - \frac{c}{s-\xi}\right) > 1 - \frac{c}{\alpha}.$$

The reservation offer for $L^m$, hereafter denoted $\alpha_m$, is again obtained by equating the LHS with the RHS of the above inequality and solving for $\alpha$:

$$\alpha_m = \frac{c}{1 - \lambda(1 - \frac{c}{s-\xi})}.$$

A wage bill-maximizing leadership, denoted $L^{wb}$, chooses to strike if $\lambda s - \xi > \alpha$, meaning that its reservation offer $\alpha_{wb}$ is given by:

$$\alpha_{wb} = \max\{\lambda s - \xi, 0\}.$$

The reservation offer for $M$, $\overline{\alpha}$, is identical to that in (2.1). Figure 2.1 plots all three reservation offers along the $\lambda$ axis, for a small value of $\xi$. It can be seen that $\alpha_{wb}$ intersects $\overline{\alpha}$ at a high level of $\lambda$ and hits zero at a low level of $\lambda$, implying that $L^{wb}$ will settle when it is either very weak ($\lambda < \hat{\lambda}_{wb}$) or powerful ($\lambda > \tilde{\lambda}_{wb}$). $L^m$, on the other hand, begins to settle at a lower level of power ($\hat{\lambda}_m$) than $\hat{\lambda}$ in (2.5). $\alpha_{wb}$ represents the expected wage premium to be had were workers to go on strike, and in the region of $\lambda$ where $L^{wb}$ strikes but $L^m$ settles ($\lambda \in (\hat{\lambda}_m, \tilde{\lambda}_{wb})$), $\alpha_{wb}$ is always above $\overline{\alpha}$, the maximum share $L^m$ can hope to win by settlement. Thus in this range of $\lambda$ the expected share of the surplus going to workers under $L^{wb}$ is strictly superior to any negotiated share $\hat{\alpha} \in (\alpha_m, \overline{\alpha})$ obtainable under $L^m$. It is also straightforward to show that the collective welfare of workers, inclusive of their utility

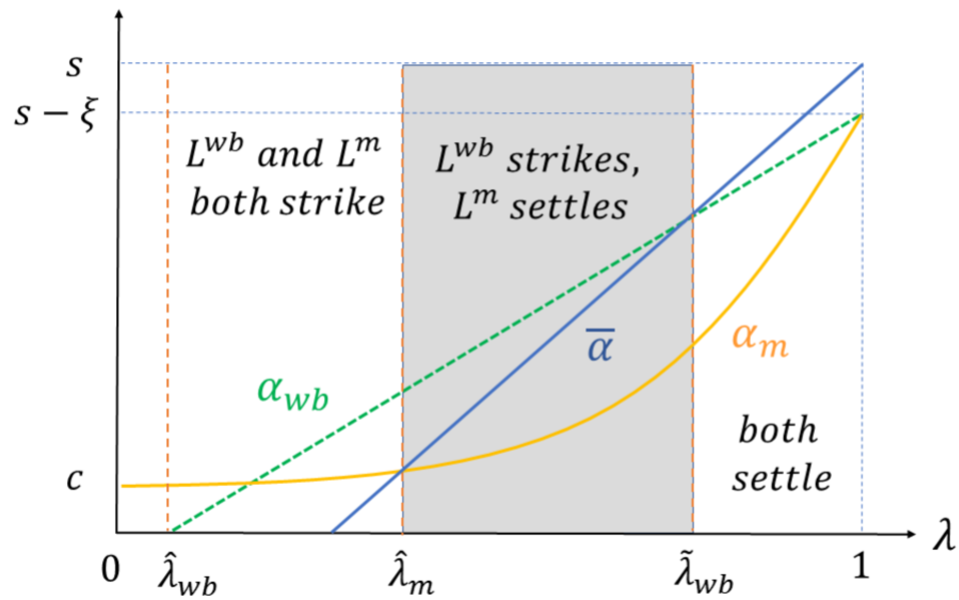



Figure 2.1: Reservation offers of $M$, $L^m$, and $L^{wb}$ in the presence of strike cost

derived from union membership, is also higher under $L^{wb}$.[2]

The result that a membership-maximizing leadership $L^m$ underserves workers when it is powerful enough mostly carries through when we use a worker utility-maximizing (instead of wage-bill maximizing) union leadership as the benchmark. A worker utility-maximizing $L$, hereafter denoted $L^u$, would choose to strike whenever:

$$\frac{\lambda}{2}(s-\xi-c)\left(1-\frac{c}{s-\xi}\right)+\lambda s-\xi>\max\left\{\frac{1}{2}(\alpha-c)(1-c/\alpha),0\right\}+\alpha, \tag{2.7}$$

for any $\alpha < s$ that can be obtained through settlement. As before, the reservation offer for

[2]The proof is as follows: recall that the utility of worker $i$ is $U_i = (i\cdot\alpha - c)\mathbb{1}\{join\}+\alpha$. In the event of a strike, $E\alpha = \lambda s - \xi$. We need to show that in the range of $\lambda$ where $L^m$ settles for $\hat{\alpha}\in(\alpha_m,\overline{\alpha})$ but $L^{wb}$ would strike, workers' expected utility from striking, $E\int_0^1 U_i|_{strike}di$, is greater than their expected utility from settling, $\int_0^1 U_i|_{settle}di$. It must be proven, in other words, that for $\lambda\in(\hat{\lambda}_m,\tilde{\lambda}_{wb})$,

$$\frac{\lambda}{2}(s-\xi-c)(1-\frac{c}{s-\xi})+\lambda s-\xi>\frac{1}{2}(\hat{\alpha}-c)(1-\frac{c}{\hat{\alpha}})+\hat{\alpha}.$$

Since we know that $\lambda s-\xi>\hat{\alpha}$ in the relevant range of $\lambda$, a sufficient condition for the above inequality is $\lambda(s-\xi-c)(1-\frac{c}{s-\xi})>(\hat{\alpha}-c)(1-c/\hat{\alpha})$. Since $s-\xi>\hat{\alpha}$, it follows that $1-\frac{c}{s-\xi}>1-c/\hat{\alpha}$. It suffices to show, then that $\lambda(s-\xi-c)>\hat{\alpha}-c$, which is easy to do: $\lambda s-\xi-c+(1-\lambda)(\xi+c)>\lambda-\xi-c>\hat{\alpha}-c$.

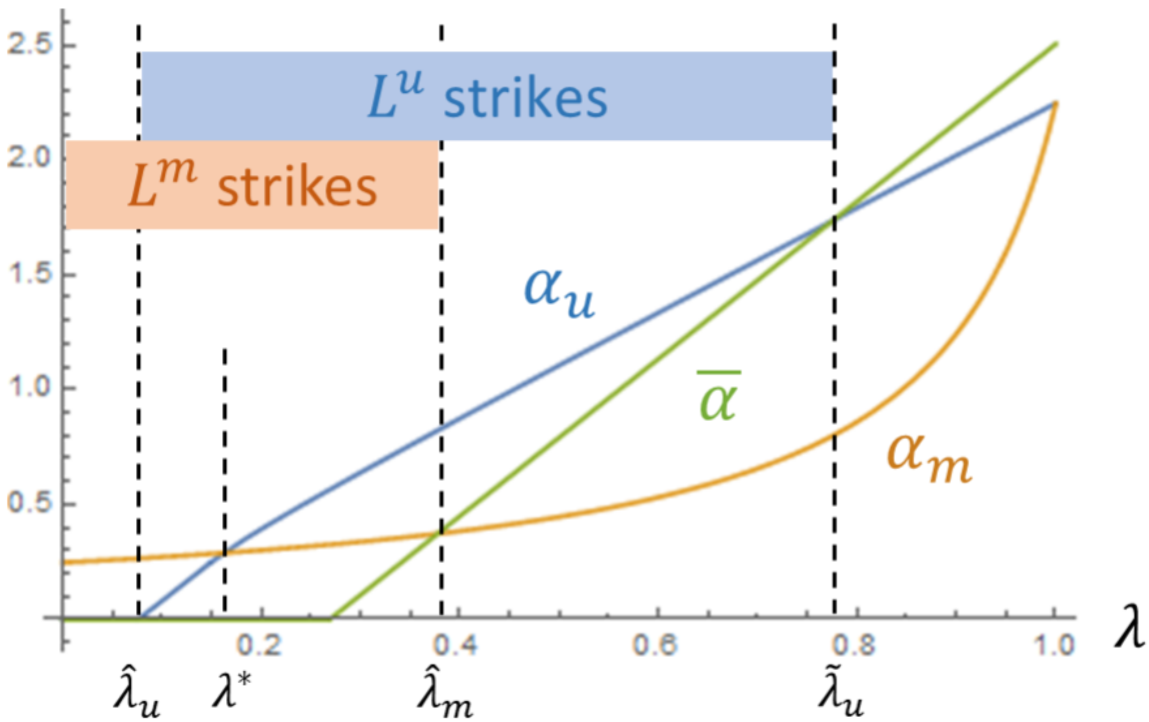
$L^u$, denoted $\alpha_u$, is found by solving for $\alpha$ that turns (2.7) into an equality.

$\alpha_u$ is a piecewise function with two thresholds in $\lambda$. The first, $\hat{\lambda}_u$, is the $\lambda$ value that makes the LHS of (2.7) equal to zero: below $\hat{\lambda}_u$, striking results in negative expected utility for workers and therefore is worse than accepting a share of zero without a fight. Thus at this range of union power $L^u$'s reservation offer is zero. The second threshold, $\tilde{\lambda}_u$, is that which turns $\alpha_u$ equal to $c$. $\alpha$ below $c$ is a share that is low enough to ensure a membership of zero: therefore the part of worker utility that derives from union membership ($\frac{1}{2}(\alpha - c)(1 - c/\alpha)$) no longer enters $L$'s calculus when considering the benefits of settling; all it cares about is securing a wage premium that workers will get to enjoy even in the absence of the union. Thus for $\lambda < \tilde{\lambda}_u$, $\alpha_u$ is just the LHS of (2.7). To summarise, $\alpha_u$ as a function of $\lambda$ takes the following form:

$$\alpha_u = \begin{cases} 0 & \text{if } \lambda < \hat{\lambda}_u \\ \frac{\lambda}{2}(s-\xi-c)(1-\frac{c}{s-\xi}) + \lambda s - \xi & \text{if } \hat{\lambda}_u \leq \lambda < \tilde{\lambda}_u \\ \frac{1}{6}\Big[\lambda(3s-\xi-2c) - 2(\xi+c) & \\ +\sqrt{(\lambda(3s-\xi-2c)-2(\xi+c))^2 - 12c^2}\Big] & \text{if } \tilde{\lambda}_u \leq \lambda \end{cases}$$

where $\hat{\lambda}_u = \frac{2\xi(s-\xi)}{3s^2+\xi^2+c^2-2c(s-\xi)-4s\xi}$ and $\tilde{\lambda}_u = \frac{2(\xi+c)(s-\xi)}{3s^2+\xi^2+c^2-2c(s-\xi)-4s\xi}$. Figure 2.2 plots $\alpha_u$, $\alpha_m$, and $\overline{\alpha}$ as functions of $\lambda$. The $\alpha_u(\lambda)$ curve intersects with the $\alpha_m(\lambda)$ curve at two points: first, at $\lambda = \frac{2(\xi+c)(s-\xi)}{3s^2+\xi^2+c^2-2c(2s-\xi)-4s\xi} \equiv \lambda^*$, which is larger than $\tilde{\lambda}_u$; second, at $\lambda = 1$. Given the concavity of $\alpha_u(\lambda)$ and the convexity of $\alpha_m(\lambda)$, this implies that $\alpha_m(\lambda) \leq \alpha_u(\lambda)$ for all $\lambda \geq \lambda^*$. The implications for the behavioral difference between $L^m$ and $L^u$ are the following:

1. At low levels of union power, $L^m$ prefers to strike while $L^u$ settles. This is because $L^m$ never accepts any offer below $c$, whereas $L^u$ is willing to accept any offer when the chances of winning a strike is too small to justify the sacrifice it entails.




Note: graph plotted using parameter values $s = 2.5$, $c = 0.25$, $\xi = 0.26$, $b = 0.9$

Figure 2.2: Reservation offers of $M$, $L^m$, and $L^u$ in the presence of strike cost

2. At an intermediate range of $\lambda$, both types of $L$ either strike or settle, but $L^u$ never settles when $L^m$ would strike.

3. At high levels of $\lambda$, both types settle.

Simply put, $L^u$ requires more power than does $L^m$ to begin deploying the strike weapon, but also has to acquire greater levels of power before it shelves that weapon.

What are the welfare implications of their divergent behavior? To answer, we can consider four cases:

- Case 1: both $L^u$ and $L^m$ strike.
- Case 2: $L^u$ settles and $L^m$ strikes.
- Case 3: $L^u$ strikes and $L^m$ settles.
- Case 4: both settle.

It is straightforward to see that (expected) worker welfares are the same in case 1. In case 2, the fact that $L^u$ settles implies that a negotiated $\alpha$ was obtainable that makes workers weakly better off than striking, hence $L^m$ could only have harmed worker welfare by striking. In case 3, the fact that $L^u$ strikes implies that no $\alpha$ acceptable to $M$ could have made workers better off than striking, therefore whatever offer $L^m$ has accepted must have been inferior to the alternative of striking. Case 4 is more interesting. It was seen that $\alpha_u > \alpha_m$ for $\lambda \in (\lambda^*, 1)$: therefore in this upper range of $\lambda$ the bargained share $\hat{\alpha} = \beta\alpha_j + (1-\beta)\overline{\alpha}$, $j = u, m$ must be greater for $L^u$ than $L^m$, assuming bargaining power $\beta$ to be the same for both. Worker utility will be higher under $L^u$ as a result. But in the lower range $\lambda < \lambda^*$, $L^m$ has the higher reservation offer, and will likely secure a higher negotiated share. This is the one instance where workers may actually benefit from having a leadership that is more concerned with its own survival than with worker welfare.

To summarise the foregoing welfare analysis, we have seen that a membership-maximizing union leadership will tend to be less combative and willing to settle for less than what workers would deem justified given the union's fighting capacity. Intuitively, this is because such a union leadership does not care about the intensity of members' loyalty to the organization: its aim is to make the greatest possible number of workers just loyal enough to hold on to their membership cards. To this end, pushing a modest demand that management can tolerate is more prudent policy than risking a showdown. Workers may feel dissastisfied, even betrayed, by the timidity of the leadership, but knowing that a union is still better than no union at all, and lacking a competing union they can switch allegiance to, will feel obliged to retain their membership.

**Robustness of qualitative results to alternative modeling assumptions**

The central implication of the baseline model, namely that a membership-maximizing union bureaucracy under-represents the interests of workers when they are capable of settling without fighting, rests on two qualitative results: (i) $M$'s reservation offer $\overline{\alpha}$ is lower than workers' threat payoff, (ii) $L$'s reservation offer $\underline{\alpha}$ is even lower than $\overline{\alpha}$ for values of $\lambda$ above a certain threshold.

These results were derived using a number of simplifying assumptions that perhaps sacrificed too much descriptive realism. In this sub-subsection we will see that the qualitative results still hold under alternative, more realistic assumptions.

One possible weakness of the baseline model is the assumption that $M$'s distaste for the union's very existence ($b$) does not vary with either the strength ($\lambda$) of the union or its membership. The result is that a union organizing a small portion of the workforce and wielding only negligible power is treated as equally troublesome as a union that organizes nearly everyone and is capable of bringing management to heel. Although the scale of money and effort employers often expend to fend off unionization drives give an indication of their distaste for unions no matter their strength and size, a more realistic assumption would be that unions generally cause more trouble for management the stronger and better organized they are – even before considering the higher wage concessions they would command as a result.

To address this weakness, we begin by assuming that $b$ is an increasing function of $\lambda$ only, with $b(0) = 0$ and $b(1) \equiv \bar{b}$. Thus $M$'s payoffs under striking and bargaining are now altered in the following way:

$$U_M|_{strike} = \lambda(-b(\lambda)) + (1-\lambda)s,$$
$$U_M|_{settle} = s - \alpha - b(\lambda).$$

Equating the two payoffs and solving for $M$'s reservation offer yields:

$$\overline{\alpha}(\lambda) = \max\{\lambda s - (1-\lambda)b(\lambda), 0\}.$$

It can be seen that as long as $b(\,.\,) > 0$, $\overline{\alpha}$ is always below workers' expected share under a strike, $\lambda s$. Differentiating $\overline{\alpha}(\lambda)$ with respect to $\lambda$, we have:

$$\overline{\alpha}'(\lambda) = s - (1-\lambda)b'(\lambda) + b(\lambda).$$

Evaluating the above derivative at $\lambda = 1$, we have that $\overline{\alpha}'(1) = s + \bar{b}$. It can be checked that the derivative with respect to $\lambda$ of $L$'s reservation offer, $\underline{\alpha}$ in (2.2), evaluates to $\frac{s}{c}(s-c)$ when $\lambda = 1$. For small enough $c$ (or large enough $\bar{b}$), it can be seen that $\underline{\alpha}'(1) > \overline{\alpha}'(1)$. Therefore the $\underline{\alpha}(\lambda)$ curve would lie below the $\overline{\alpha}(\lambda)$ curve in the range of $\lambda$ above some threshold (call it $\hat{\lambda}'$) where the two curves intersect. So for all $\lambda \in [\hat{\lambda}', 1]$, $L$ and $M$ would settle and the resulting wage premium for workers would be smaller than their expected premium from striking by at least $(1-\lambda)b(\lambda)$. Thus the qualitative result that powerful enough unions under-represent the interest of workers is preserved.

Next, let us consider the case where $M$'s distaste depends only on the size (membership) of the union. Let $b(m)$ be strictly increasing in $m = 1 - \frac{c}{\alpha}$, with $b(0) = 0$ and $b(1 - \frac{c}{s}) = \bar{b}$. Then $M$'s payoffs are:

$$U_M|_{strike} = \lambda(-\bar{b}) + (1 - \lambda)s,$$
$$U_M|_{settle} = s - \alpha - b(\lambda).$$

$\overline{\alpha}$ is defined implicitly by the expression:

$$\overline{\alpha} = \lambda(s + \bar{b}) - b\left(1 - \frac{c}{\overline{\alpha}}\right) \quad \text{if } \overline{\alpha} > c.$$

It can be checked that $\overline{\alpha} = s$ when $\lambda = 1$. Treating $\overline{\alpha}$ as a function of $\lambda$ and totally differentiating both sides with respect to $\lambda$, we find after rearranging:

$$\overline{\alpha}'(\lambda) = \frac{s + \bar{b}}{1 + b'(1 - c/\overline{\alpha})c/\overline{\alpha}^2}.$$

This derivative evaluates to $\frac{s+\bar{b}}{1+b'(1-c/s)c/s^2}$ at $\lambda = 1$. If $\frac{\bar{b}}{b'(1-c/s)} > \frac{c}{s}$, the slope of $\overline{\alpha}(\lambda)$ at $\lambda = 1$ is steeper than the slope of $\lambda s$ (workers' expected share of $s$ under a strike). Again, for small enough $c$ (or large enough $\bar{b}$), it can be checked that the slope of $\underline{\alpha}$ $(= \frac{c(1-c/s)}{(1-\lambda(1-c/s))^2})$ is steeper than that of $\overline{\alpha}$ at $\lambda = 1$; therefore the $\underline{\alpha}(\lambda)$ curve would lie below the $\overline{\alpha}(\lambda)$ curve in the range of $\lambda$ above some threshold where the two curves cross. By the same reasoning as above, this implies the qualitative results of the baseline model carries through in this case as well.

Another potential pitfall of the baseline model lies in the assumption that unions completely dissolve upon losing a strike. While it is certainly reasonable to suppose unions that repeatedly lose would eventually cease to exist, in a one-period model such elimination may be too extreme an assumption. Despite having lost a strike and therefore winning nothing, some workers may still choose to remain union members out of a sense of duty and/or anticipation that the union could do better next time, and so the union may yet live.

In this case, if we assume as in the baseline model that $M$'s distaste for the union $b$ is a fixed constant irrespective of the size of the union, then $b$ becomes irrelevant to $M$'s decision on whether to fight or settle (since it cannot get rid of the union either way). It follows that $M$'s reservation offer in this case would equal $\overline{\alpha} = \lambda s$, so that a bargained settlement ($\hat{\alpha} \leq \overline{\alpha}$) no longer *necessarily* disappoints workers.

But what if we allowed $b$ to vary with $m$ as we have previously done? Suppose the function $b(m)$ is as defined above. To operationalize the assumption of positive union membership even after winning zero concession, we redefine $m$ to equal $1 - \frac{c}{\gamma+\alpha}$, where $\gamma > c$: it is as though workers believe the union benefits their colleagues by a money equivalent of $\gamma$ just by virtue of existing (e.g. by offering an 'employee voice' option rather than just the 'exit' option). Let $\underline{m} = 1 - \frac{c}{\gamma}$ and $\overline{m} = 1 - \frac{c}{\gamma+s}$, and define $\underline{b} = b(\underline{m})$ and $\overline{b} = b(\overline{m})$. Then $M$'s payoffs under striking and bargaining become:

$$U_M|_{strike} = \lambda(-\overline{b}) + (1-\lambda)(s - \underline{b}),$$
$$U_M|_{settle} = s - \alpha - b\left(1 - \frac{c}{\alpha}\right).$$

Solving for $\overline{\alpha}$ yields:

$$\overline{\alpha}(\lambda) = \lambda(s + \overline{b} - \underline{b}) + \underline{b} - b\left(1 - \frac{c}{\gamma + \overline{\alpha}}\right),$$

which implicitly defines $\overline{\alpha}$ as a function of the parameters. It is easily checked that $\overline{\alpha}(0) = 0$

and $\overline{\alpha}(1) = s$. Totally differentiating both sides with respect to $\lambda$ and rearranging, we find:

$$\overline{\alpha}'(\lambda) = \frac{s + \bar{b} - \underline{b}}{1 + b'\left(1 - \frac{c}{\gamma + \overline{\alpha}}\right) \cdot \frac{c}{(\gamma + \overline{\alpha})^2}}.$$

This derivative, evaluated at $\lambda = 1$, is equal to $\frac{s + \bar{b} - \underline{b}}{1 + b'(\overline{m})c/(\gamma + s)^2}$. This slope would exceed $s$ if we assume high enough $\bar{b} - \underline{b}$ and small enough $c$ and $b'(\overline{m})$, which would mean that the $\overline{\alpha}$ curve lies beneath the $\lambda s$ curve for some interval of $\lambda$ below 1. If, moreover, the $\underline{\alpha}$ curve lies below $\overline{\alpha}$ over a similar interval of $\lambda$, the qualitative results of the baseline model would survive.

To derive $\underline{\alpha}$, note that $L$'s payoffs under bargaining and striking have also changed:

$$U_L|_{strike} = \lambda(1 - \frac{c}{\gamma + s}) + (1 - \lambda)(1 - \frac{c}{\gamma}),$$
$$U_L|_{settle} = 1 - \frac{c}{\gamma + \alpha}.$$

Equating the two payoffs and solving for $\alpha$ yields:

$$\underline{\alpha}(\lambda) = \frac{\lambda \gamma s}{s + \gamma - \lambda s}.$$

It follows that $\underline{\alpha}(0) = 0$, $\underline{\alpha}(1) = s$, $\underline{\alpha}'(\lambda) = \frac{\gamma s(s+\gamma)}{(s+\gamma-\lambda s)^2}$, $\underline{\alpha}'(1) = \frac{s(s+\gamma)}{\gamma}$, and $\underline{\alpha}'(0) = \frac{\gamma s}{s+\gamma}$.

Comparing these to the corresponding derivatives of $\overline{\alpha}$, We can see that for sufficiently small $\gamma$ (i.e. small enough $\underline{m}$), the slope of $\underline{\alpha}$ will be steeper than that of $\overline{\alpha}$ in a neighborhood of $\lambda = 1$; hence $\overline{\alpha} > \underline{\alpha}$ in that neighborhood (possibly even for all $\lambda \in (0, 1)$). Thus in this case, too, $L$ ends up agreeing to a settlement that is inferior to fighting when $L$ is powerful enough (or possibly even when it is not, i.e. always).

Lastly, it may be asked whether a strike would still be more beneficial to workers when the leadership would rather settle, if there is more than one period on the horizon. Given that a strike entails the risk of the union getting destroyed, and therefore of workers not getting anything in future periods, intuition might suggest that a strike-averse (or compromise-prone) union leadership actually works to the benefit of workers. We shall see that this is not the case.

Assume the same game described in the baseline model is repeated over two periods. If $L$ loses a strike in the first period, it gets eliminated and there is no union to represent workers in the second period – workers in both periods therefore get nothing. Otherwise, the union makes it to the beginning of the second period and inherits $\lambda$ from the first period. We retain the simpler assumption in the baseline model that $b$ is invariant to union power or membership. At period 1 all agents discount payoffs in the next period by a common factor $\delta$.

Exogenous parameters are fixed in both periods, but endogenous variables may differ across periods. We therefore index endogenous outcomes with the subscript $t$, so that $\hat{\lambda}_2$ refers to the threshold of union power that permits peaceful settlement in period 2, $\overline{\alpha}_1$ denotes $M$'s reservation offer in period 1, etc.

All outcomes in the second (and final) period are identical to those in the baseline model. So given $\lambda$ in the first period, agents anticipate that conditional on the union surviving, workers' (expected) second-period share of the surplus will be $\hat{\alpha}_2(\lambda)$ (the negotiated wage premium as in (2.3)) if $\lambda \geq \hat{\lambda}_2$ (the threshold in (2.5)) and $\lambda s$ if $\lambda < \hat{\lambda}_2$.

These second-period outcomes affect $L$ and $M$'s decision-making in the first period. Workers, however, are assumed to behave as if they are not forward-looking: they make decisions to join the union solely on the basis of concessions won in the current period (i.e. $m_t = 1 - \frac{c}{\alpha_t}$. I make this assumption for the sake of mathematical tractability: the introduction of forward-looking workers may alter the results to some extent.

To solve the game, we need to consider two cases: $\lambda \geq \hat{\lambda}_2$ and $\lambda < \hat{\lambda}_2$. Let us begin with the former. $M$'s lifetime utilities if it takes a strike or settles in $t=1$ are the following:

$$U_M|_{strike} = \lambda(-b+\delta(s-\hat{\alpha}_2(\lambda)-b)) + (1-\lambda)(s+\delta s),$$
$$U_M|_{settle} = s-\alpha_1-b+\delta(s-\hat{\alpha}_2(\lambda)-b).$$

Equating and solving for $M$'s reservation offer in $t=1$ yields:

$$\overline{\alpha}_1(\lambda) = \max\Big\{\underbrace{\lambda s-(1-\lambda)b}_{=\overline{\alpha}_2}-\delta(1-\lambda)(\hat{\alpha}_2(\lambda)+b),\, 0\Big\}.$$

It can be seen that $\overline{\alpha}_1$ is lower than $\overline{\alpha}_2$. This is intuitive: a strike in $t=1$ affords $M$ an opportunity to rid itself of the union not only in the current period but also for the remaining period. The shadow of the future makes fighting that much more attractive, and so $M$ is less willing to pay for peace.

As for $L$, lifetime utilities are:

$$U_L|_{strike} = \lambda\Big(1-\frac{c}{s}+\delta\big(1-\frac{c}{\hat{\alpha}_2(\lambda)}\big)\Big)+(1-\lambda)\cdot 0,$$
$$U_L|_{settle} = 1-\frac{c}{\alpha_1}+\delta\Big(1-\frac{c}{\hat{\alpha}_2(\lambda)}\Big).$$

It follows that $L$'s $t=1$ reservation offer is:

$$\underline{\alpha}_1(\lambda) = \frac{c}{1-\lambda(1-c/s)+\delta(1-\lambda)\Big(1-\frac{c}{\hat{\alpha}_2(\lambda)}\Big)} \leq \underline{\alpha}_2 = \frac{c}{1-\lambda(1-c/s)}.$$

$L$'s reservation offer is lower than in $t=2$ for the opposite reason that $M$'s is lower: $L$ now has more to lose from an unsuccessful strike than just its current-period membership, therefore $L$ has a greater incentive to sue for peace.

Recalling that $\hat{\alpha}_2(1)=s$, it is easy to see that $\overline{\alpha}_1(1)=\underline{\alpha}_1(1)=s$. It can also be shown that the derivatives of the $t=1$ reservation offers evaluated at $\lambda=1$ are $\underline{\alpha}_1'(1)=(1+\delta)(s-c)s/c$

and $\overline{\alpha}_1'(1) = (1+\delta)(s+b)$. For small enough $c$, this means $\overline{\alpha}_1 > \underline{\alpha}_1$ for all $\lambda \in (\hat{\lambda}_1, 1)$. In other words, $M$ and $L$ will settle in $t = 1$ whenever the fighting capacity of the union is above some threshold $\hat{\lambda}_1$ that is defined implicitly by the equation:

$$\hat{\lambda}_1 = \frac{(1+\delta)s(b+c+\delta(b-c+\hat{\alpha}_2(\hat{\lambda}_1) - bc/\hat{\alpha}_2(\hat{\lambda}_1))}{(s+b+\delta(\hat{\alpha}_2(\hat{\lambda}_1)+b))\left(s-c+\delta s(1-c/\hat{\alpha}_2(\hat{\lambda}_1))\right)}.$$

In the other case of $\lambda < \hat{\lambda}_2$, the players' payoffs are slightly altered in view of the fact that they anticipate a strike in $t = 2$:

$$\begin{aligned}
U_M|_{strike} &= \lambda\left[-b+\delta(\lambda(-b)+(1-\lambda)s)\right] + (1-\lambda)(1+\delta)s, \\
U_M|_{settle} &= s - \alpha_1 - b + \delta(\lambda(-b)+(1-\lambda)s), \\
U_L|_{strike} &= \lambda\left(1-\frac{c}{s}+\delta\Big(1-\frac{c}{s}\Big)\right) + (1-\lambda)\cdot 0, \\
U_L|_{settle} &= 1 - \frac{c}{\alpha_1} + \delta\lambda\Big(1-\frac{c}{s}\Big).
\end{aligned}$$

Reservation offers change accordingly:

$$\begin{aligned}
\overline{\alpha}_1(\lambda) &= \max\Big\{\underbrace{\lambda s-(1-\lambda)b}_{=\overline{\alpha}_2} - \delta\lambda(1-\lambda)(s+b),\, 0\Big\}, \\
\underline{\alpha}_1(\lambda) &= \frac{c}{1-\lambda(1-c/s)+\delta\lambda(1-\lambda)(1-c/s)} \leq \underline{\alpha}_2 = \frac{c}{1-\lambda(1-c/s)}.
\end{aligned}$$

Since $\overline{\alpha}_1(0) = 0$ and $\underline{\alpha}_1(0) = c$, settlement cannot happen until $\lambda$ reaches a threshold $\tilde{\lambda}$ given by:

$$\tilde{\lambda} = \frac{1}{2\delta}\left[\sqrt{(1-\delta)^2 + \frac{4\delta s(b+c)}{(s-c)(s+b)}} - (1-\delta)\right].$$

It can be shown that $\tilde{\lambda} > \hat{\lambda}_2$. Thus when $\lambda \in (0, \hat{\lambda}_2)$, $\underline{\alpha}_1$ is always above $\overline{\alpha}_1$ and settlement is impossible. Using the fact that $\hat{\alpha}_2(\hat{\lambda}_2) = \frac{c(s+b)}{s-c}$, it can be checked that $\overline{\alpha}_1$ and $\underline{\alpha}_1$ are both continuous at $\lambda = \hat{\lambda}_2 = \frac{s(b+c)}{(s-c)(s+b)}$. The implication is that $\hat{\lambda}_1$ at which the reservation offers intersect is somewhere above $\hat{\lambda}_2$. Thus a strike occurs for all $\lambda < \hat{\lambda}_1$; otherwise the parties settle.

Compared to striking, is settlement ever beneficial to workers? It suffices to consider the case of $\lambda \geq \hat{\lambda}_2$. The expected present value of workers' lifetime wage bill under striking and settling in $t = 1$ are respectively $\lambda(s + \delta\hat{\alpha}_2(\lambda))$ and $\alpha_1 + \delta\hat{\alpha}_2(\lambda)$. It follows that the $t = 1$ wage settlement $\alpha_1$ which would make workers indifferent between settling and striking equals $\lambda s - \delta(1-\lambda)\hat{\alpha}_2(\lambda)$. We have seen, however, that the maximum $M$ would be willing to concede is $\lambda s - \delta(1-\lambda)\hat{\alpha}_2(\lambda) - (1-\lambda)(1+\delta)b$. The negotiated wage premium $\hat{\alpha}_1 \in [\underline{\alpha}_1, \overline{\alpha}_1]$ will therefore give workers at least $(1-\lambda)(1+\delta)b$ less than the discounted lifetime earning workers expect to win by striking.

Compared to the baseline model (which is a special case of the current model with $\delta = 0$), we see that conditional on $\lambda \geq \hat{\lambda}_1$, workers' wage penalty from having a membership-maximizing leadership becomes larger in a dynamic setting, and moreover worsens as the future grows in importance. But because the threshold $\hat{\lambda}_1$ is higher than $\hat{\lambda}_2$, we also see that the range of $\lambda$ that leads to disappointing compromise diminishes in a dynamic setting.

The intuition behind these contrasts is the following: in the one-period model it was seen that $M$'s reservation offer $(\lambda s - (1-\lambda)b)$ was lower than workers' threat payoff by an amount equal to the expected efficiency gain from a strike, $(1-\lambda)b$. When there are multiple periods on the horizon, the efficiency gain that can result from $M$ defeating a strike in the current period extends to all remaining periods, hence the difference between $M$'s reservation offer and workers' threat payoff gets magnified by the present-discounted number of periods. In the meantime, $L$'s reservation offer also drops when there are multiple periods, but not nearly as much as $M$'s does, since $L$ is constrained by the need to win at least $c$ to survive in any period. Hence room for compromise diminishes.

We have seen in this sub-subsection that the qualitative results of the baseline model are

Table 2.1: Comparative statics of the baseline model

| | parameters | | | |
|---|---|---|---|---|
| outcome variables | $\lambda$ | $s$ | $b$ | $c$ |
| workers' bargained share of surplus | (+) | (+) | (−) | (+) |
| membership | (+) | (+) | (−) | (−) |
| strike frequency | N/A | (−) | (+) | (+) |

robust to modifying the assumptions of the model in various ways to enhance their realism. As long as a union's existence creates large enough inefficiencies that are increasing in the union's size or power, and a strike affords an opportunity to reduce or eliminate such inefficiencies, any compromise between management and a membership-maximizing union leadership would give workers less than their threat payoff. We may thus treat the predictions of the baseline model as the main takeaways from this subsection.

**Comparative statics**

We now turn to making predictions about how changes in the parameters of the baseline model may affect various collective bargaining outcomes. To keep the analysis tractable, we shall ignore the strike cost parameter introduced in the welfare analysis and stick to the simpler model developed up to (2.5). We focus on three outcomes of interest: bargained worker share of the surplus ($\hat{\alpha}$), (expected) membership ($m$), and strike frequency per union. This last outcome is the prevalence of strikes among unionized firms, and depends on the cumulative distribution function $G(.)$ of $\lambda$ and the threshold $\hat{\lambda}$ in the unionized sector: assuming that all unions below $\hat{\lambda}$ will strike, strike frequency in the unionized sector would equal $G(\hat{\lambda})$. Differentiating the expressions for the three outcomes of interest with respect to $\lambda$, $s$, $b$, and $c$ yields the set of comparative statics presented in Table 2.1.

That $\hat{\alpha}$ and $m$ should both increase with $s$ is intuitive: with more surplus to be distributed, $L$ would be able to secure higher wages and enjoy greater popularity. The result that strike frequency declines with $s$ is apparently at odds with the well-documented procyclicality of strike activity (see Kennan, 1986, for a summary of the evidence). However, the model can be reconciled with the observed pattern if we interpret it as a model of union leaders' intention rather than as a model of strike activity in general: strikes in this model only happen when $L$ and $M$ intend to undertake a serious strike that only ends with total defeat for one side. Many strikes in practice do not go to such extremes. During economic booms (when $s$ is high), union leaders may tend to engage in token strikes under pressure from the rank and file: Ashenfelter and Johnson (1969)'s classic model of strikes under asymmetric information directly addresses this possibility. Low unemployment may also imbue rank-and-file militants with enough confidence to undertake strikes independently of the leadership, such as in wildcat strikes. Hence strike activity measured in counts can increase during economic upswings while life-or-death struggles decline in frequency. Indeed, studies that find strike duration (i.e. intensity) to be counter-cyclical (Vroman, 1989) do lend support for the prediction of my model that serious strikes become more frequent during economically hard times.

$\hat{\alpha}$ and $m$ are both decreasing in $b$ because as long as $L$ wishes to avoid a strike, it needs to compensate an employer that has a greater level of distaste for the union with a lower wage premium; membership drops as a result. A strike is more likely the more the employer dislikes the union, since room for peaceful compromise is diminished.

An increase in $c$ has the effect of raising $\hat{\alpha}$ because $L$ is incentivized to deliver more for the workers in order to persuade them to join despite the higher cost. But even a higher negotiated wage is not enough to keep some members at the margin from deserting. This is because $L$'s reservation offer $\underline{\alpha}$ is set to equate the membership resulting from settlement to expected membership under a strike, and the latter declines with higher $c$: thus by construction, $\underline{\alpha}$ doesn't rise enough to counter the effect of a rising $c$ on $m$, and neither does $\hat{\alpha} = \beta\overline{\alpha} + (1 - \beta)\underline{\alpha}$, unless $L$'s bargaining power ($\beta$) also rises. As $L$'s reservation offer increases with $c$, room for peaceful compromise diminishes and strike frequency rises.

We now summarize the key results from this subsection into the following proposition:

**Proposition 2.1.** *A membership-maximizing union leadership $L^m$ exhibits the following behavior patterns:*

*(1) $L^m$ wishes to go on strike only if $\lambda$ (its probability of winning) is sufficiently low; otherwise it prefers to settle.*
*(2) Conditional on settling, accepts a wage offer that is weakly inferior to what workers can expect to win by striking.*
*(3) Compared to a worker utility-maximizing leadership $L^u$, generally achieves inferior outcomes for workers, except in the special case when $\lambda$ is low enough and both $L^m$ and $L^u$ are willing to settle.*

### 2.3.2 The possibility of co-optation by management

The baseline model did not require $M$ to play any active role in taming the union bureaucracy: all that was required for $L$ to underrepresent workers was for $M$ to dislike the union. I now consider a more ominous possibility, where $M$ has the power to manipulate $c$ and make it conditional on $L$'s behavior.

Let $d$ denote a subsidy that $M$ can offer to $L$ to help reduce the cost of membership, and let $k$ denote the default cost of membership in the absence of this subsidy. Hence $c = k - d$. The subsidy can include the provision of company property for use as union office space, allowing union activity during company time, payment of part of the salaries of union officials (which need not amount to bribery in a technical sense)[3], or agreeing to deduct 'agency fees' from non-member paychecks to cover $L$'s bargaining costs (as is commonly done in non-'right-to-work' states in the U.S.). Such forms of subsidy can do one of two things: they reduce

[3]such practice is prohibited in principle in the U.S. under Section 8(a)(2) of the National Labor Relations Act, but Section 302(c) of the Labor Management Relations Act allows for exceptions.

the opportunity cost of joining a union, or they help reduce the cost (borne by workers) of maintaining a union, and allows $L$ to charge lower membership fees. We assume $d$ is costly to $M$; we also restrict $d$ to be strictly less than $k$ ($d \geq k$ would represent a case of a company-dominated union, outlawed under the NLRA). Importantly, $d$ is provided only on condition that the workers' share of the surplus negotiated be no more than a certain cutoff: $\alpha \leq \epsilon$. Then if $L$ ever participates in this deal, it would obviously choose $\alpha = \epsilon$: And so, depending on whether $L$ chooses to participate, the utilities of $L$ and $M$ would be the following:

$$EU_L = \begin{cases} 1 - \frac{k-d}{\epsilon} & \text{if participate} \\ 1 - \frac{k}{\hat{\alpha}} & \text{if not participate and settle} \\ \lambda(1 - \frac{k}{s}) & \text{if strike} \end{cases}$$

$$EU_M = \begin{cases} s - b - \epsilon - d & \text{if } L \text{ participates} \\ s - b - \hat{\alpha} & \text{if } L \text{ does not participate and settles} \\ s - \lambda(s + b) & \text{if take a strike} \end{cases}$$

Now suppose the fighting capacity of the union is above the threshold in (2.5) required to settle without fighting ($\lambda \geq \hat{\lambda}$). Then the participation constraint for $L$ is:

$$1 - \frac{k-d}{\epsilon} \geq 1 - \frac{k}{\hat{\alpha}(\lambda)} \iff \epsilon \geq \frac{k-d}{k}\hat{\alpha}(\lambda). \tag{2.8}$$

For $M$, the benefits of offering this carrot has to be higher than the alternative, so $M$'s participation constraint is:

$$s - b - \epsilon - d \geq s - b - \hat{\alpha}(\lambda) \iff \hat{\alpha}(\lambda) - \epsilon \geq d. \tag{2.9}$$

Since $M$ holds the initiative, it will make (2.8) bind by setting $\epsilon = \frac{k-d}{k}\hat{\alpha}(\lambda)$. Thus $M$'s problem becomes:

$$\max_d \left\{ s - b - \frac{k-d}{k}\hat{\alpha}(\lambda) - d = s - b - \hat{\alpha}(\lambda) + \Big(\frac{\hat{\alpha}(\lambda)}{k} - 1\Big)d \right\}.$$

Since $\frac{\hat{\alpha}(\lambda)}{k} > 1$ by assumption, $M$'s solution to the problem is to offer the most generous subsidy possible under the constraints $d < k$ and (2.9). Let this optimal subsidy be $d^* = k - \delta$, where $\delta > 0$ is an arbitrarily small quantity. Then the optimal cutoff for $\alpha$, $\epsilon^* = \frac{\delta}{k}\hat{\alpha}(\lambda)$ can also be made arbitrarily small. These solutions satisfy (2.9) since

$$\hat{\alpha}(\lambda) - \epsilon^* - d^* = \hat{\alpha}(\lambda)\Big(1 - \frac{\delta}{k}\Big) - (k - \delta)\Big(\frac{\hat{\alpha}(\lambda)}{k} - 1\Big) > 0$$

$$\Longleftrightarrow \hat{\alpha}(\lambda) - \epsilon^* \geq d^*.$$

This result raises the possibility of almost complete co-optation by management of unions that are powerful enough to settle without a strike: membership in such unions would be the same as in non-co-opted counterparts, but their leadership would deliver virtually nothing for their members. Only a serious electoral challenge to incumbent leadership may be able to compel a change in $L$'s behavior in such cases.

Next we consider the case of a union with $\lambda < \hat{\lambda}$. The participation constraints for $L$ and $M$ now become:

$$\text{For L: } 1 - \frac{k-d}{\epsilon} \geq \lambda\Big(1 - \frac{k}{s}\Big)/, \Longleftrightarrow \epsilon \geq \frac{k-d}{1 - \lambda(1 - k/s)}.$$

$$\text{For M: } s - b - \epsilon - d \geq s - \lambda(s+b) \Longleftrightarrow \lambda s - (1-\lambda)b - \epsilon \geq d.$$

Following the same reasoning as before, we find that $M$'s optimal offer $(\epsilon^*, d^*)$ consists of $\epsilon^* = \frac{\delta}{1-\lambda(1-k/s)}$ and $d^* = k - \delta$. For arbitrarily small $\delta$ and $\epsilon^*$, $M$'s participation constraint is approximately $\lambda s - (1-\lambda)b \geq k$, which can be rewritten as:

$$\lambda \geq \frac{b+k}{s+b}.$$

Therefore $M$ would only offer the conditional subsidy to unions above this critical fighting capacity. Weaker unions will not even be considered worthy of co-optation.

The possibility analyzed in this subsection is admittedly an extreme one. Nevertheless, it illustrates the way in which the problems of union bureaucracy can be exacerbated in practice by ostensibly benign managerial efforts to accommodate labor leaders.

Union security agreements that grant some measure of financial security to unions, such as automatic check-off of dues, may be thought of as a type of company subsidy: the company helps to alleviate union officers' burden of collecting dues, at some administrative cost to itself. Reflecting on the "collaborationist" character of labor organizations in the 1970s, Lynd and Lynd (2014) wrote that unions "have become a new kind of company union, financially independent of the rank and file because the company deducts union dues from the worker's pay check." Similarly, Moody (2010) observed that automatic dues check-off, a demand won through militant struggles, had the unintended consequence of accelerating the trend toward bureaucratization by further insulating union officials from control by the membership.

A less direct example of how managerial accommodation can entail a Faustian bargain between organizational growth and quality of representation comes from the Service Employees International Union (SEIU). Under the presidency of Andrew Stern in the 1990s and 2000s, SEIU became one of the largest and fastest-growing unions in the U.S. One controversial aspect of SEIU's growth was Stern's strategy of 'bargaining to organize', which involved 'neutrality agreements' between the union and firms targeted for unionization. These agreements committed SEIU to cooperate with management in raising productivity and lobbying for greater governmental subsidies for their industry, in return for employers' promise not to interfere with unionization drives (Lichtenstein, 2013). In line with their business-friendly strategy, SEIU agreed to numerous concessions including wage freezes, benefit cuts, and removal of job security, at times waging war against more militant locals that insisted on better deals (Winslow, 2010a).

The key results from this subsection can be summarized as follows:

**Proposition 2.2.** *When the employer $M$ can manipulate the cost of union membership, the following outcomes ensue:*

*(1) $M$ offers to make union membership practically costless, on condition that $L^m$ accepts a share for workers $\hat{\alpha}$ that is practically zero, if $\lambda$ is high enough. Otherwise, $M$ makes no such offer and takes a strike.*
*(2) Conditional on such an offer being made, $L^m$ accepts and retains the same membership despite $\hat{\alpha}$ being arbitrarily small.*

### 2.3.3 The dynamic evolution of union strength and membership

Up to this point I have treated $\lambda$ as an exogenous quantity, which facilitated a comparative static analysis of membership determination. But it is reasonable to suppose that membership, thus determined, would in turn have a feedback effect on $\lambda$ in subsequent periods. This suggests a process of dynamic co-evolution between union density and union strength, an exploration of which may help us better understand the decline of unions in the U.S. and elsewhere over the last few decades.

We will thus allow $\lambda$ at the beginning of a given period t + 1 to be a function of existing membership attained in the previous period, $m_t$, in order to explore potential pathways of union membership growth or decay over time. We will also introduce a technology parameter $T$ as an argument for $\lambda$ (so that $\lambda_{t+1} = \lambda(m_t, T)$, where $T$ can be interpreted as the degree of ease with which a given share of workers can successfully disrupt production in the event of a strike, owing to technical conditions such as the scarcity of skills employed or the intricacy of the division of labor at the firm. We assume throughout that $\lambda(.,.)$ is monotonically increasing in both arguments.

For the moment, let us treat $\lambda(.)$ as a univariate function of $m_t$. If initial membership is such that $\lambda(m_t)$ is above the threshold $\hat{\lambda}$ in (2.5), then $L$ would choose to bargain and the resulting membership $m_{t+1}$ will be given by:

$$m_{t+1}(m_t) = 1 - \frac{c}{\hat{\alpha}(\lambda(m_t))}. \tag{2.10}$$

If, on the other hand, $\lambda(m_t) < \hat{\lambda}$, a strike occurs and with probability $\lambda(m_t)$ the union wins all of $s$; the resulting member share would equal $1 - c/s$. This would represent the initial member share for all unions that has just recently won a strike. Then equation (2.10) applies to such unions for membership in the subsequent period.

Let $\underline{m} = \lambda^{-1}(\hat{\lambda})$ denote the minimum membership required to ensure a peaceful settlement. A steady-state member share $m^{ss} \in [\underline{m}, 1 - c/s]$ will then be implicitly defined by the following equality:

$$\begin{aligned} m^{ss} = m_{t+1}(m^{ss}) &= 1 - \frac{c}{\hat{\alpha}(\lambda(m^{ss}))} \\ &= 1 - \frac{c}{\beta\left[\lambda(m^{ss})s - (1-\lambda(m^{ss}))b\right] + (1-\beta)\left[\frac{c}{1-\lambda(m^{ss}(1-c/s)}\right]}. \end{aligned} \tag{2.11}$$

$m^{ss}$ is stable if the derivative of the membership function evaluated at $m^{ss}$ is smaller than 1 in absolute value:

$$m'_{t+1}(m^{ss}) = c \cdot \frac{\partial\hat{\alpha}/\partial\lambda}{\hat{\alpha}^2}\lambda'(m^{ss})$$
$$= \frac{c\left[\beta(s+b) + (1-\beta)\frac{c(1-c/s)}{(1-\lambda(m^{ss})(1-c/s))^2}\right]}{\left[\beta\left[\lambda(m^{ss})s - (1-\lambda(m^{ss}))b\right] + (1-\beta)\left[\frac{c}{1-\lambda(m^{ss})(1-c/s)}\right]\right]^2}\lambda'(m^{ss}) < 1.$$

It can be seen that this condition would hold for small enough $\lambda'(m^{ss})$. However, if $\lambda'(m^{ss})$ is large enough, then $m^{ss}$ is not stable and a union with initial membership below $m^{ss}$ will tend to bleed members over time, while those with $m_0 > m^{ss}$ will see membership continually grow until hitting some upper bound $\overline{m} = 1 - c/(\hat{\alpha}(\overline{\lambda}))$, where $\overline{\lambda}$ is an arbitrary upper bound of the range of $\lambda(.)$ assumed to be below 1 (because in practice no union completely dominates an employer, in the sense of being able to win a strike with certainty).

Based on the foregoing analysis, we may discern the following four categories of unions with distinct membership dynamics. The shapes of possible membership functions for each category are illustrated in Figure 2.3.

1. Fragile unions: unions for which $m^{ss}$ does not exist in the interval $[\underline{m}, 1-c/s]$ and $m_{t+1}(m_t) < m_t$ for all $m_t > \underline{m}$. These unions face a perpetual tendency for membership to decline, and will be characterized by frequent and periodic strikes motivated by a need to reverse the tide of declining membership.

2. Unstable unions: unions for which $m'_{t+1}(m^{ss}) > 1$ for $m^{ss} \in (\underline{m}, \overline{m})$. The fate of such a union diverges sharply depending on whether its initial membership exceed the threshold of $m^{ss}$: those with $m_t < m^{ss}$ suffer the same problem as fragile unions, until they win a strike; those with $m_t \geq m^{ss}$ become more securely established, with membership eventually reaching $\overline{m}$. They remain vulnerable, however, to membership shocks that can push $m$ below $m^{ss}$.

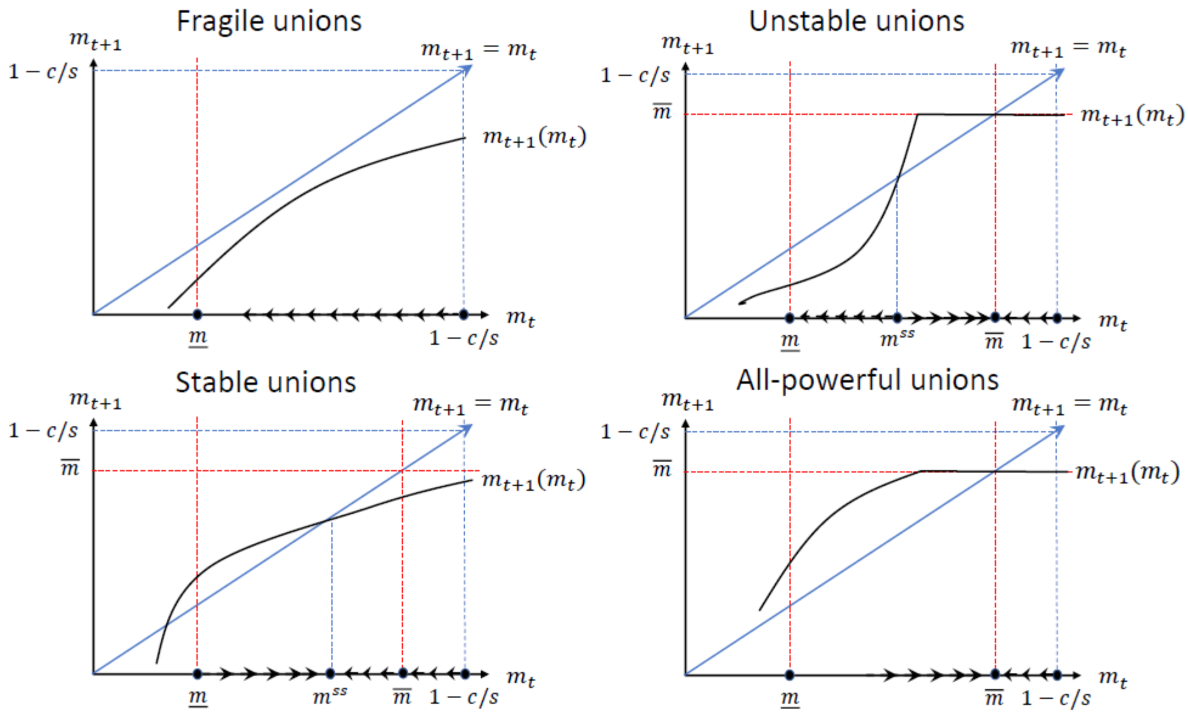



Figure 2.3: Four categories of unions with distinct membership dynamics

3. Stable unions: unions for which $m'_{t+1}(m^{ss}) < 1$, for $m^{ss} \in (\underline{m}, \overline{m})$. For these unions, both the share of the surplus going to workers (i.e. wage premia) and membership would tend to converge to an intermediate level. Strikes would rarely be used, except possibly at the time of birth of a union.

4. All-powerful unions: unions for which $m^{ss}$ does not exist in the interval $[\underline{m}, \overline{m})$, and $m_{t+1}(m_t) \geq m_t$ for all $m_t > \underline{m}$. Such unions enjoy an exceptionally high and enduring level of membership and share of the surplus going to workers. Of all the categories of unions, they are the most resilient to membership shocks. Rarely would they resort to strikes.

Which of the four categories a union falls into depends on the slope of its $\lambda$ function over the domain of $m_t$. This slope, in turn, may depend on the technical conditions prevailing at the particular workplace which makes shutting down production easier or harder to achieve with a given share of union members. This would therefore be an appropriate place to introduce the technology parameter $T > 0$ as an argument of the $\lambda$ function. For concreteness, we define $\lambda(m_t, T)$ explicitly, using the simplest possible specification:

$$\lambda(m_t, T) = \min\{Tm_t, 1\},$$

Since $\lambda_{t+1}$ has to be at least $\hat{\lambda} = \frac{b+c}{(s+b)(s-c)}$ for peaceful settlement to occur in period $t+1$, the minimum member share in period $t$ that is necessary for peaceful settlement in the next period is given by:

$$\underline{m} = \frac{\hat{\lambda}}{T} = \frac{(b+c)s}{(s+b)(s-c)T}.$$

For all $m_t \geq \underline{m}$, then, the membership function is given by:

$$m_{t+1}(m_t) = 1 - \frac{c}{\beta(Tm_t s - (1 - Tm_t)b) + (1-\beta)\left(\frac{c}{1-Tm_t(1-c/s)}\right)}.$$

To facilitate exposition, we focus on two extreme cases: $\beta = 1$ and $\beta = 0$. We begin with the case where $L$ has all the bargaining power ($\beta = 1$). The membership function then simplifies to:

$$m_{t+1}(m_t) = 1 - \frac{c}{Tm_t(s+b) - b},$$

and

$$m'_{t+1}(m_t) = \frac{c(s+b)T}{(Tm_t(s+b) - b)^2}.$$

It can be seen that for $m_t > \underline{m}$, $m'_{t+1}(m_t)$ is uniformly declining in $m_t$, implying that if $m^{ss} \in (\underline{m}, \overline{m})$ exists, it will constitute a stable steady-state. Solving for $m^{ss}$ in the equation

$m(m^{ss}) = m^{ss}$ yields the following steady-state membership equation:

$$m^{ss} = \frac{1}{2} + \frac{b + \sqrt{(b + bT + sT)^2 - 4T(b + c)(b + s)}}{2T(b + s)} < \overline{m} = 1 - \frac{c}{s}.$$

A real-valued $m^{ss}$ exists if $T \geq \frac{b+2c+2\sqrt{c(b+c)}}{b+s}$. It can be shown that this same conditions guarantees that $\partial m^{ss}/\partial T > 0$. Thus unions in firms with $T < \frac{b+2c+2\sqrt{c(b+c)}}{b+s}$ cannot attain a stable membership level, and would fall in the category of fragile unions. Those with $T$ above this threshold and with $m^{ss} \in (\underline{m}, \overline{m})$ will fall in the category of stable unions. As $T$ increases, there eventually comes a point where $m^{ss} = \overline{m}$, at which point the union becomes all-powerful.

We now repeat the same analysis for the other extreme case: $\beta = 0$. Then $m_{t+1}(m_t)$ simplifies to:

$$m_{t+1}(m_t) = Tm_t(1 - c/s).$$

The slope of $m_{t+1}(.)$ then equals $T(1-c/s)$ over the entire range of $m_t$. Thus the $m_{t+1}$ curve never crosses the 45 degree line in the $(m_t, m_{t+1})$ plane except in the special case of $T = \frac{s}{s-c}$. Either $m_{t+1}$ is always above the 45-degree line (if $T > \frac{s}{s-c}$) or below (if $T < \frac{s}{s-c}$). In other words, a union would either become fragile or all-powerful depending on the magnitude of $T$.

This simple model illustrates the importance of technical conditions of production in shaping union power in the long run. Standardization of tasks and the accompanying degradation of skills, for instance, can erode the collective action capacity and thereby the strength of unions by making each would-be striker or unionist more easily replaceable (see Montgomery, 1980 for a classic account of how the 'scientific management' movement of the early $20^{\text{th}}$ century undermined the crafts unions in the U.S.). Technical innovations such as conveyor belts or Just-In-Time (JIT) production, which render the entire production chain more

Table 2.2: Comparative statics of steady-state union membership

| | parameters | | | |
|---|---|---|---|---|
| outcome | $b$ | $c$ | $s$ | $T$ |
| $m^{ss}$ | $(-)$ | $(-)$ | $(+)$ | $(+)$ |

vulnerable to disruptions at any single point, may work in the opposite direction: indeed, Silver (2003) explains the traditional strength of unions in the transportation and some manufacturing sectors precisely by their 'workplace bargaining power' owed to the workers' ability to effectively disrupt production.

We may thus predict that union density and union wage premia would both be greater in high-$T$ sectors. The analysis in Section 2.2.1 implies that high-$T$ sectors would also be the ones where the problems of union bureaucracy would be more marked, and where officials of existing unions would be less willing to organize strikes. This does not necessarily imply that unions in high-$T$ sectors should be less strike-prone, as rank-and-file pressure on the leadership to mount strikes could also be greater in such sectors.

We conclude this section by deriving comparative statics of $m^{ss}$ with respect to various parameters. Totally differentiating (2.11) with respect to $b$, $c$, $s$, and $T$ and rearranging yields the results in Table 2.2, all of which are intuitive.

## 2.4 Conclusions

Union officials throughout history have betrayed the wishes of those they represent often enough to be perceived by many labor movement practitioners and intellectuals as constituting a bureaucracy distinctly removed from the rank and file. One explanation for their behavior may simply be that lay workers are susceptible to irrational outbursts of militancy, and that union leaders serve as a rational check against the herd mentality of their followers. But if we accept the assumption of human rationality that underpins neoclassical economics, we must seriously consider the possibility that the leadership vs. rank and file divide arises from a real conflict of interests.

The model of union bureaucracy developed in this chapter demonstrates that when union leaders maximize membership rather than some measure of worker welfare, they can under-represent the interests of the workers once they become powerful enough. This result is driven by the assumption that a union, by its very existence, imposes a cost to the employer. The employer is therefore tempted to try to eliminate the union by decisively defeating it in open battle (i.e. a strike/lockout), should the opportunity arise; a peaceful settlement is only attractive to the employer if the union can compensate the employer for tolerating its existence. Leaders of unions that are strong enough to compensate their employers in this manner and still deliver high enough wages to retain members are then compelled to accept a compromise rather than engage in battle, even if that compromise leaves workers with a smaller share than what they can win in expectation by unleashing their full fighting capacity. This is because for such a union leadership, any increase in membership that may result from a victorious strike appears marginal relative to the loss of members that can result from defeat. In a sense, concern with membership maximization makes union officials more risk-averse than the rank and file with regard to strikes.

In an extension of the baseline model, it was seen that the tendency for the union bureaucracy to under-serve workers can be exacerbated when management is able to co-opt the leadership by offering to subsidize their member recruitment and retention effort. This points to the corrupting potential of union security agreements that are usually considered 'labor-friendly'.

A limitation of the model is that it is built on rather extreme assumptions about strikes, which are treated as winner-take-all contests with life-or-death consequences. Future work must extend the model to accommodate less extreme varieties of strikes, such as short strikes intended to 'blow off steam' or to build members' loyalty to the organization.

Another limitation of the model is that it assumes workers have no real choice over whom to elect as their leaders. Allowing for contested leadership elections may still yield the same qualitative results if any faction that rises to the leadership has a long-term interest in preserving the union apparatus (despite the short-term risk of being voted out if they compromise on disappointing terms), but the seriousness of the leadership vs. rank and file divide would most likely be mitigated.

# Chapter 3

# What Do Right-to-Work Laws Do to Unions? Evidence from Six Recently Enacted RTW Laws

## 3.1 Introduction

The decline of unions in the U.S. over the past half-century has been attributed to many possible causes. Alongside globalization, decline of manufacturing, and Reaganism, an oft-cited culprit is the migration of union jobs from the Midwest to the relatively union-free South (Lichtenstein, 2013; Moody, 2014; Friedman, 2008). The weakness of unions in the U.S. South, in turn, is commonly attributed to so-called right-to-work (RTW) laws adopted by most southern states in the early post-war years: these prohibit unions from collecting mandatory fees from non-members covered by collective bargaining agreements, thereby increasing incentives for workers to free-ride on union services. For many advocates of organized labor, the lower rates of unionization and wages in southern right-to-work states is taken as evidence of the deleterious effects right-to-work laws can have on unions and worker well-being in states that would newly adopt them.[1] Proponents of RTW laws, on the other hand, argue mandatory union dues violate the rights of individuals by forcing them to support a political cause. The U.S. Supreme Court sided with the latter view in its 2018 *Janus v AFSCME* decision, making the entire U.S. public sector right-to-work; the Protecting the Right to Organize (PRO) Act, which the House of Representatives passed in March 2021, would nullify existing RTW laws if the Senate also passes it. If opponents of RTW are right, the outcome of such legal and legislative battles will have far-reaching consequences.

Researchers have long debated whether RTW laws actually caused the observed low rates of unionization in RTW states. The debate has never been settled, largely owing to data limitations and the ensuing challenges to causal identification: most RTW laws were enacted by states with similar characteristics (including strong anti-union preferences) at a time when consistent data series on outcomes of interest were unavailable. However, the adoption of RTW laws by six U.S. states (Oklahoma, Indiana, Michigan, Wisconsin, West Virginia, and Kentucky) since the turn of the century provides an opportunity to identify RTW effects using better data and methods. For the first time in the literature, this chapter examines the impact of RTW laws passed in all six states on a range of union-related outcomes. Using a mix of empirical methods including difference-in-differences, event studies, and synthetic controls, I find evidence that in the private sector, RTW laws decrease union coverage by more than 10 percent, all else equal. Surprisingly, I find RTW laws to have only a small and insignificant effect on free-riding behavior as measured by the share of unionized workers who are nonmembers. Moreover, union formation through NLRB-administered elections do

[1] A typical argument along this line can be found in Shermer, Elizabeth Tandy (2018, April 24) *The right to work really means the right to work for less*, The Washington Post. https://www.washingtonpost.com/news/made-by-history/wp/2018/04/24/the-right-to-work-really-means-the-right-to-work-for-less/ (accessed October 12, 2020). Also see Neal, Candy (2012, January 13) *Local union opposes right-to-work movement*, The Herald https://duboiscountyherald.com/b/local-union-opposes-right-to-work-movement (accessed May 12, 2021) for a sampling of how union activists view the issue.

not appear to be adversely affected by RTW. I also find evidence that RTW legislation increases union wage differentials by up to five percentage points, which is suggestive of a change in union bargaining behavior; levels of strike activity appear to be unaffected by RTW laws. In the public sector, I find RTW laws to be associated with declines in union coverage comparable to the private sector, but it is difficult to separate the causal effect of RTW laws from the effects of policies targeting public sector unions that in many cases preceded RTW legislation. Separately evaluating the effects of *Janus v. AFSCME* on public sector unions in states that never adopted RTW laws, I find union coverage in the affected states to have changed little, and union wage differentials to have increased by more than five percent. Some of these findings are quite novel, and challenge conventional assumptions about how RTW laws impact unions.

The remainder of this chapter is organized as follows: Section 3.2 discusses the nature of RTW laws, the findings of previous studies regarding their effects, and how the more recently enacted RTW laws under study may shed some light on old questions; Section 3.3 presents my empirical strategy and data; Section 3.4 presents the results; and Section 3.5 concludes.

## 3.2 Background on Right-to-Work Laws and the Literature on their Effects

### 3.2.1 The nature of RTW laws and their controversies

Contrary to what their name implies, right-to-work laws have nothing directly to do with protecting workers' right to a job. Their chief function is to outlaw union security clauses in collectively bargained labor contracts. Union security historically referred to agreements requiring employees to either join the union as a condition of employment (as in closed-shop and union-shop agreements) or to pay a fee equivalent or commensurate to membership dues even if they choose not to join – these are called agency-shop agreements, and the fees known as agency fees or 'fair share dues' (see Blair, 1974 for a summary of various types of union security clauses). Union security was devised as a remedy for the free-rider problem inherent in the U.S. system of collective bargaining that took root in the New Deal period: a union that is recognized (through a formal election) as the exclusive bargaining agent for eligible employees in a bargaining unit is required to represent all such employees, regardless of their membership status. Hence absent union security, workers would have an incentive to benefit from union representation without paying their fair share of its cost[2]. The Taft-Hartley Act (Labor Management Relations Act) of 1947 made mandatory union membership illegal, and gave state legislatures discretion to prohibit weaker forms of union security in their jurisdictions[3]. Thus the subsequently enacted right-to-work laws essentially sought to eliminate agency-shops and agency fees (which can range from 20 to 80 percent of full membership dues [Swindle, 1984]). Ultimately 27 of the 50 U.S. states came to adopt RTW laws (as of 2021), the great majority of them located in the U.S. South.

Because of the obvious possibility of free-riding that RTW laws would create (and the perceived consequent threat to union membership and finances), organized labor has strongly opposed RTW laws, characterizing them as union-busting tools that would ultimately hurt all workers by depriving them of opportunities for union representation. Their proponents (usually backed by business interests) have challenged the constitutionality of agency fees on the grounds that they violate the freedom of individuals not to support a political cause they do not believe in (see Cantor, 1983 for a discussion of legal controversies surrounding agency fees). There is likely to be some truth in this assertion, given the substantial financial

[2]This cost is non-trivial: as documented by Bennett (1991), annual dues and fees collected by unions amounted to $504 (current dollars) per union member in 1987

[3]Taft-Hartley did allow union-shop agreements (unless banned by RTW laws), which technically did require new hires to become union members within some prescribed period as a condition of employment, but a court decision in the 1951 *Union Starch Refining v. National Labor Relations Board* case made union-shop clauses practically unenforceable (Cogen, 1954).

and organizational contributions of unions to Democratic election campaigns (Bennett, 1991; Dark, 2000; Lichtenstein, 2013). The U.S. Supreme Court sided with the proponents' view in its 2018 *Janus v AFSCME* decision, making agency fees illegal in the entire U.S. public sector. A more recently developed argument in favor of RTW is that it benefits the economy of the adopting state by attracting new investment (Devinatz, 2015).

### 3.2.2 Older literature on RTW laws

The critics' claim that RTW laws snuff out unionism is made credible by the fact that union density tends to be lower in RTW states; moreover, the share of nonmembers among unionized workers (who would constitute free-riders under RTW) is markedly higher in RTW states compared to non-RTW states – 17.3 percent against 7.4 percent by Sobel (1995)'s estimate. For researchers, however, establishing the causal nature of this relationship has been a persistent challenge. 20 out of the 21 states that were RTW at the end of the 20$^{th}$ Century had adopted the laws prior to 1980, when consistent data series on union membership was unavailable. Moreover, the predominantly southern RTW states were more hostile to unions compared to the rest of the nation even before the 1947 Taft-Hartley Act, in part because unionism threatened to subvert the long-standing racial hierarchy in the South (Farhang and Katznelson, 2005; Dixon, 2007). It was therefore very difficult for earlier studies to separate any causal effect of RTW laws from the effect of pre-existing anti-union preferences that may have led those states to enact RTW laws in the first place (see Moore and Newman, 1985 and Moore, 1998 for a comprehensive review). Findings based on analyses of cross-sectional data have been mixed: Studies that use aggregate data at the state level and properly address simultaneity and omitted variables issues tend to find no independent effect of RTW laws on union density (Lumsden and Petersen, 1975; Hunt and White, 1983; Moore et al., 1986; Farber, 1984); studies using individual- or department-level data tend to find significantly negative effects (Hundley, 1988; Ichniowski and Zax, 1991;

Davis and Huston, 1995).

In theory, RTW laws can reduce union density (share of union members among all employees) through two channels. The 'Free-rider Hypothesis' holds that the free-rider problem makes unions costlier to organize and maintain, so that the supply of union services would decline in the long run; the 'bargaining power' hypothesis holds that the decrease in membership caused by free-riding weakens the bargaining power of unions, which lowers the benefits of, and hence the demand for, unions. Under both scenarios, union coverage (share of workers covered by union contracts) declines as the result of an increase in the share of non-members (free-riders) among unionized workers[4]: the effect on union density would thus represent the composite effect of the increase in free-riding and the resulting decline in coverage.

Not many studies have attempted to separately identify RTW effects on free-riding and union coverage, however. To the extent of my knowledge, only three papers have tried to estimate effects on free-riding: Davis and Huston (1993) find RTW laws to increase the share of free-riders by around 8 percentage points; focusing on public sector bargaining laws, Hundley (1993) estimates that rules that permit agency shops has little effect on share of covered nonmembers; Sobel (1995) estimates the extent of 'true free-riding' to be quite small and concludes that repealing RTW laws would increase union membership only slightly. Thus the central mechanism by which RTW laws are alleged to stifle unionism has rarely been empirically tested. As for union coverage, Ellwood and Fine (1987) use annual NRLB election data going back to 1950, and find that RTW laws roughly halve the annual flow of newly unionized workers into the stock of the already unionized, in the first five years after their initial enactment (the effect dissipates after a decade or so). The authors note that this could be attributed either to the Bargaining Power Hypothesis or to the symbolic/psychological impact of a high-profile legislative defeat for organized labor. Apart

[4]The share of nonmembers does not *necessarily* have to rise to bring about this outcome, if the share of nonmembers is large to begin with: the loss of agency fee revenue from existing nonmembers alone could be enough to reduce the supply and demand for unions. But this is not very likely given the low percentage of covered nonmembers in non-RTW states.

from union membership, researchers have looked at RTW effects on wages and employment. Here again, findings are mixed: Moore (1980), Wessels (1981), and Moore et al. (1986) find negative but insignificant effects on wages, while Reed (2003) finds a significant positive effect after controlling for initial economic conditions in RTW states; Holmes (1998) finds a large increase in manufacturing employment share when crossing the border from a RTW state to a non-RTW state, but Stevans (2009) finds no significant RTW effect on employment growth or business relocation after controlling for states' general business climates; Kalenkoski and Lacombe (2008) only find RTW laws to be associated with a two percent increase in the share of manufacturing employment. Again, lack of data going sufficiently back in time made it impossible to credibly isolate the effect of RTW laws.

### 3.2.3 RTW legislation in the 21$^{st}$ Century

Beginning in the early 2000s, six U.S. states newly passed RTW laws: Oklahoma (2001), Indiana (2012), Michigan (2012), Wisconsin (2015), West Virginia (2016), and Kentucky (2017). The circumstances surrounding passage of these laws were similar in many respects: all were introduced by Republican-dominated legislatures, all were advertised by proponents as a way to attract businesses and promote job growth, and all were met by union-mobilized protests numbering in the thousands (Devinatz, 2015). Their language is also very similar to one another, and have been interpreted by courts to apply equally to the private and public sectors (Feigenbaum et al., 2018). None were retroactively applicable to labor contracts signed before the laws took effect.

The context of their passage differed in other respects. Oklahoma was doing rather well economically in the run-up to 2001 thanks to the oil and gas boom (Eren and Ozbeklik, 2016), whereas the other states were still reeling from the aftermath of the 2009 recession; Oklahoma's RTW law, unlike the others, was put to a referendum in the form of an amendment to the state constitution. RTW laws in Indiana, Michigan, and Wisconsin have their shared origins in the sweeping Republican victory in the November 2010 mid-term elections, which gave Republicans monopoly control over 11 state governments. According to Lafer (2013), the Republicans swept to power in those states – as well as the corporate interests backing them – were eager to seize upon the rare but possibly temporary opportunity to advance legislative goals that had long lingered on wish lists, RTW being one of them. Thus in 2011-12, seventeen other state legislatures introduced RTW bills alongside Indiana and Michigan; in 2013 and 2014 the number grew to 21 and 20 state legislatures. Devinatz (2015) hypothesizes that beyond economics motives, RTW laws may have been aimed at removing a key source of organizational support for Democrats that labor unions were deemed to be. Thus it is possible that IN, MI, and WI enacted RTW as part of a nation-wide effort to entrench Republican power rather than in response to specific local conditions.[5] West Virginia's case is a bit different in that the newly Republican-controlled state legislature had to override the then-Democratic Governor's veto. WV is also unique in that a series of court injunctions against enforcement of its RTW law (pending a legal challenge mounted by some unions) delayed its implementation by over a year since initial passage.[6] Kentucky's RTW law came two month after Republicans had won control of the state's lower house for the first time since 1921. The only other state to pass a RTW law in this period was Missouri in 2017, but the law was repealed by referendum in the following year and never took effect.

These recently enacted RTW laws provide an opportunity to estimate their impacts more credibly using panel data that are presently available. Not many studies have examined the effects of these recent RTW laws, however. Eren and Ozbeklik (2016) focus on Oklahoma using the synthetic control method and find some negative effect (on the order of 20

[5]Lafer (2013) makes a similar claim regarding many anti-labor policies pursued by the same cohort of Republican state governments, such as restrictions on public sector collective bargaining – they bore no relation to local conditions such as levels of state and local government debt

[6]Redmond, Sean P. (2020, April 27) *Twice as nice, Court upholds West Virginia right-to-work law for good*, U.S. Chamber of Commerce. `https://www.uschamber.com/article/twice-nice-court-upholds-west-virginia-right-work-law-good` (accessed May 11, 2021).

percent) on union density and no effect on employment; Manzo and Bruno (2017) estimate a difference-in-differences model using RTW adoption in three midwestern states (IN, MI, and WI) and find negative effects on unionization and average wages. Makridis (2019) use RTW legislation in four of the six switching states (OK and KY excluded) to identify their effect on the subjective well-being of union members, and surprisingly find RTW laws to increase union members' satisfaction on the job: the author conjectures that this may be driven by increased pressure on unions to deliver quality services to the employees they seek to recruit. Lapmann (2015) makes a similar diagnosis based on their finding that Indiana's RTW law had, if anything, a positive impact on unions and unionized workers. Feigenbaum et al. (2018) use RTW legislation in five of the six states (KY excluded) to estimate their effect on Democratic voter turnouts, and find that RTW laws lead to a 2 to 3.5 percentage fall in Democratic vote shares (which the authors attribute to reduced union capacity for Democratic campaign contributions).

What the present chapter does, for the first time in the literature, is to examine the impact of RTW laws in all six recent-adopting states, on a range of union- and labor market-related outcome variables: union coverage, share of covered nonmembers, strike activity, union wage differential, and union and non-union wages. The next section presents the details.

## 3.3 Data and Methodology

The availability of consistent data series on union membership and other outcomes of interest before and after passage of RTW laws in the six late-adopting states allows me to employ a difference-in-differences (DID) research design to identify the effects of RTW laws on various outcomes. This involves estimating a DID equation of the following form:

$$y_{st} = a_s + b_t + \beta RTW_{st} + X\Gamma + \epsilon_{st}, \tag{3.1}$$

where $s$ indexes states and $t$ indexes year. The outcomes of interest, $y$, are the following:

- Union coverage, or share of workers covered by union contracts ('coverage'), and log of union coverage ('lcoverage')
- Share of free-riders, i.e. of non-union members among covered (unionized) workers ('freeride').[7]
- Inflow of newly unionized workers per thousand non-union workers ('inflow_perK'), Number of unions newly formed through NLRB certification elections ('cert'), the resulting inflow of newly unionized workers ('inflow' for absolute numbers, 'inflow_perK' for number newly unionized per thousand non-union workers'), the number of unions exiting through NLRB de-certification elections ('decert') and the resulting outflow of unionized workers ('outflow' and 'outflow_perK').
- Strike activity as measured in annual number of strike participants per thousand unionized workers ('strikehazard'), and average strike duration measured in days ('duration').
- State-year-specific union wage differential ('udiff') conditional on observable individual characteristics. The procedure for constructing this variable is described in Appendix B.1.
- The mean log hourly wages of unionized workers ('uwage') and of non-union workers ('nuwage').

[7]Note that such nonmembers would only be free-riders in the true sense under an RTW regime; in non-RTW states they are at worst 'cheap riders' given that they do pay agency fees, which are usually cheaper than full membership dues (Swindle, 1984).

$X_{st}$ is a vector of state-level time-varying controls that may include:

- Macroeconomic variables such as log of state per-capita GDP ('lgdp'), their annual growth rates ('Dgdp'), employment rate ('emp'), and unemployment rate ('unemp'). These are likely to affect union density, bargaining power, and behavior independently of RTW laws. Also included are share of manufacturing employment ('manuf'), as union membership tends to be concentrated in this sector, and share of the labor force who have college degrees or higher ('college'), which would affect productivity and wages.

- Proxies of state political and ideological climate such as whether a Democratic state governor has been in office for most of the year ('dgov'), the share of Democratic votes in the last general election to elect the state's lower house ('house_dem'), whether Republicans or Democrats controlled both legislative chambers ('R_legctrl' and 'D_legctrl'), whether Republicans or Democrats had monopoly control over the state government ('R_trifecta' and 'D_trifecta'), and measures of citizen and government ideology ('ctz_ideo' and 'gvt_ideo') developed by Berry et al. (1998). Inclusion of these variables may help to net out the confounding effects of a broader change in the policy environment and political attitudes, correlated with RTW adoption, that may be hostile to organized labor.

Note that these variables are potentially bad controls (especially democratic vote shares, in light of Feigenbaum et al., 2018), as they may be themselves affected by RTW legislation. I will return to this issue later.

$a_s$ and $b_t$ are respectively state- and year- fixed effects. RTW is a dummy variable that switches on for the six switching states, initially in the year in which the state has been exposed to the law for at least three months. The initial year of exposure is 2001 for Oklahoma (its RTW law took effect in September 2001), 2012 for Indiana (effective as of

March 2012), 2013 for Michigan (enacted December 2012, took effect March 2013), 2015 for Wisconsin (March 2015), 2017 for West Virginia (enacted February 2016 but effective as of September 2017), and 2017 for Kentucky (January 2017). Thus $\beta$ is our coefficient of interest that identifies the effect of RTW laws on the outcome variable.

The expected sign of $\beta$ for each of the outcome variables are as follows. Under the Free-rider Hypothesis and the Bargaining Power Hypothesis, we would expect RTW laws to reduce the stock of union coverage as well as the flow of newly unionized workers, while increasing the share of covered nonmembers. So $\beta$ would be negative for 'coverage' and 'inflow', and positive for 'freeride'. This last hypothesis, in particular, will be tested using panel data for the first time in the literature.

The Bargaining Power hypothesis implies that $\beta$ would be negative for the union wage differential and union wages. But it is also conceivable that a union made more insecure by RTW would be pressured to fight more aggressively to deliver gains for workers, in an attempt to boost workers' loyalty and discourage free-riding. This is what my baseline model in Chapter 2 (as well as its extension featuring union security agreements) predicts. Makridis (2019) applies the same reasoning when interpreting their result that union members' satisfaction on the job improves under RTW. One labor law scholar (Mitchell, 1978) made the following observation:

> A union worried about re-election (or financial stability) generally becomes more difficult in its demands, more entrenched at the bargaining table, and more vociferous and unyielding in grievance processing. In fact, an insecure union has been described by one commentator as "paranoid" and "pseudo-militant".

If such 'pseudo-militant' behavior pays off, the union wage differential and union wages could actually increase. Farber (1984) offered a different reason union wage differentials could rise under RTW: attrition of unions in sectors with low benefits to unionization could pull up the average differential among the remaining unions. Thus the expected sign of $\beta$ is ambiguous for 'udiff' and 'uwage'. The association between RTW laws and union wage differentials has been examined by Moore (1980), Farber (1984), and Hundley (1993); all find union wage differentials to be higher in RTW states, but only Farber (1984) finds a significant difference.

If indeed RTW induce unions to become more assertive in their demands, then unless employers become more accommodating we would expect labor strife as expressed in strikes to become more frequent. To the best of my knowledge, this hypothesis has only been empirically tested by Wessels (1981), who finds no effect of RTW laws on strike frequency, and by Gramm (1986), who finds RTW laws to be associated with greater strike propensity and severity. Gramm (1986)'s interpretation of this result is that RTW laws, by discouraging union membership, lowers participation in strikes (assuming non-members are less likely to participate) and thereby make employers more willing to take a strike. My own model in Chapter 2 makes a similar prediction with regard to a weakening in the collective action capacity of a union. It is also possible, however, that unions with a winnowed membership and depleted strike funds may think twice before launching a strike.

Finally, the expected sign of $\beta$ is uncertain when the outcome is non-union wages ('nuwage'): it would be positive if the Republicans are right and negative if union activists are right about the consequences of RTW for jobs and labor standards.

A limitation of the DID model in (3.1) is that it does not allow the effect of the intervention to vary with time. However, given that all the RTW laws under examination applied to collective bargaining agreements coming up for renewal after the laws took effect – instead of applying retroactively to all existing agreements – and given that a typical contract is renewed every 32.8 months by one estimate (Murphy, 1992), it would be reasonable to suppose that it would take at least three years for a RTW law to fully 'bite'. In the meantime, its impact may only have been felt gradually. To allow for time-varying effects, I will also use the more flexible 'event-study' approach. Where necessary, I will augment the analysis by applying the synthetic control method due to Abadie et al. (2010), which extends the DID approach to allow comparison between units exhibiting differential trends.

My sample covers the period from 1994 to 2019. The choice of this rather long time frame is motivated by a substantial variation in the timing of RTW legislation among the treated states, and by a need to secure sufficiently many pre-periods for the earliest-treated state (Oklahoma, in 2001). I use data from the U.S. Bureau of Labor Statistics (BLS) and Federal Mediation and Conciliation Service (FMCS) to construct the strike activity measures; the U.S. Current Population Survey (CPS, Outgoing Rotation Group Files) to construct union coverage, share of free-riders, union wage differentials, union and non-union wages, employment and unemployment rates, share of manufacturing employment, and share of college graduates; U.S. Bureau of Economic Analysis (BEA) data for state per-capita GDP. Data on state government partisan control were constructed from tables published online by the National Council of State Legislatures, and from Klarner (2018)'s State Legislative Election Returns dataset. Measures of citizen and government ideology are from Berry et al. (1998), updated by the authors to cover the years up to 2016 (for citizen ideology) and 2017 (for government); for the remaining years in my panel for which these measures are missing, they are estimated using other variables. Data on NLRB union elections come from Schaller (2019)'s compilation of NLRB election records up to 2017. The detailed procedure for constructing each variable is explained in a data appendix to this chapter (Appendix B.1).

I estimate RTW effects separately for outcomes in the private sector and in the public sector. The rationale is that industrial relations in the two sectors are governed by very different legal systems: the National Labor Relations Act apply uniformly to private-sector industrial relations in all U.S. states (RTW laws being the only source of between-state variation), whereas public-sector industrial relations are governed by a highly diverse patchwork of state- and local-government laws and ordinances (Lichtenstein, 2013). Unions embedded in such disparate institutional settings may not respond to RTW laws in the same way.

Moreover, some of the RTW laws were preceded by major state-level reforms that restricted the bargaining rights of public sector unions (Lafer, 2013), which potentially confounds RTW effects in the public sector. In section 3.3 below I present and discuss results for the private sector, which I treat as my main results. In Section 3.4 I repeat the same analysis for the public sector, noting some differences to the private sector results.

## 3.4 Private Sector Results

The idea behind a difference-in-differences (DID) research design is to compare the evolution of outcomes for some 'treated' units to the evolution of the same outcomes for some 'untreated' units, or control group. I define the control group in my sample as those states that never adopted an RTW law before or during the sample period. There are a total of 24 such non-RTW states including the District of Columbia. The treatment group are the six states that switch to RTW during the sample period. Table 3.1 presents the means and standard deviations of all variables for the control states ('Never RTW'), the treated states ('Switchers'), and those that became RTW prior to 1994 ('Always RTW'). It can be seen that the treatment states are on average the most economically depressed of all groups, the most reliant on manufacturing employment (although this does not pertain to OK and WV), the least educated, the most heavily unionized, the most strike-prone, and less Democrat-leaning than the control states but not nearly as Republican as the Always-RTW states. Except for GDP per capita, share of manufacturing employment, and share of college graduates, differences between the control and treatment states tend to be slight.

Table 3.1: Summary statistics, private Sector

| | Never RTW | Switchers | Always RTW |
|---|---|---|---|

| | | | |
|---|---|---|---|
| gdp_percap | 65.531 | 48.014 | 53.002 |
| | (31.436) | (11.937) | (15.208) |
| Dgdp | 0.034 | 0.031 | 0.032 |
| | (0.029) | (0.029) | (0.035) |
| emp | 0.695 | 0.669 | 0.695 |
| | (0.037) | (0.055) | (0.051) |
| unemp | 0.057 | 0.057 | 0.052 |
| | (0.019) | (0.020) | (0.020) |
| college | 0.332 | 0.251 | 0.263 |
| | (0.077) | (0.042) | (0.050) |
| manuf | 0.104 | 0.155 | 0.115 |
| | (0.048) | (0.050) | (0.047) |
| coverage | 0.097 | 0.101 | 0.053 |
| | (0.035) | (0.037) | (0.027) |
| freeride | 0.092 | 0.079 | 0.163 |
| | (0.047) | (0.052) | (0.071) |
| inflow_perK | 1.802 | 1.664 | 0.960 |
| | (1.547) | (1.662) | (1.122) |
| outflow_perK | 0.052 | 0.049 | 0.030 |
| | (0.127) | (0.091) | (0.107) |
| uwage | 20.986 | 19.111 | 19.515 |
| | (4.732) | (3.817) | (4.597) |
| nuwage | 18.417 | 15.535 | 15.564 |
| | (4.843) | (3.408) | (3.768) |
| udiff | 0.180 | 0.210 | 0.206 |
| | (0.058) | (0.051) | (0.066) |

| | | | |
|---|---|---|---|
| strikehazard | 8.180 | 8.542 | 5.357 |
| | (19.632) | (16.851) | (15.430) |
| duration | 29.063 | 34.447 | 39.185 |
| | (29.030) | (32.194) | (46.035) |
| house_dem | 0.555 | 0.494 | 0.421 |
| | (0.154) | (0.142) | (0.168) |
| dgov | 0.547 | 0.500 | 0.275 |
| | (0.498) | (0.502) | (0.447) |
| R_legctrl | 0.244 | 0.359 | 0.608 |
| | (0.430) | (0.481) | (0.489) |
| D_legctrl | 0.545 | 0.269 | 0.234 |
| | (0.498) | (0.445) | (0.424) |
| R_trifecta | 0.131 | 0.256 | 0.496 |
| | (0.338) | (0.438) | (0.500) |
| D_trifecta | 0.295 | 0.186 | 0.106 |
| | (0.456) | (0.390) | (0.308) |
| ctz_ideo | 60.100 | 45.112 | 41.045 |
| | (13.010) | (13.796) | (9.668) |
| gvt_ideo | 54.158 | 43.112 | 37.136 |
| | (12.061) | (13.801) | (13.355) |

Means and standard deviations (SD in parentheses).

Note: gdp_percap is in thousands of U.S. dollars,

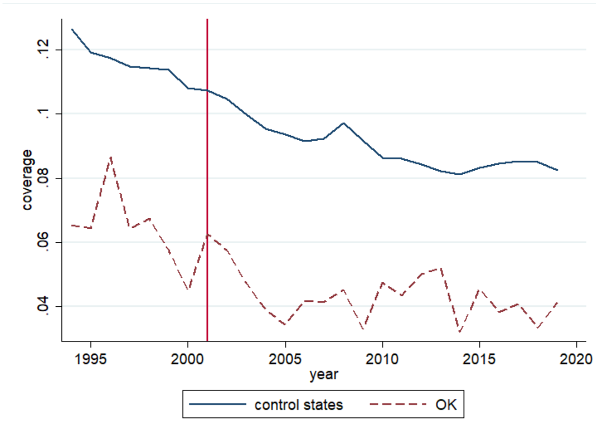
uwage and nuwage are in dollars (all nominal).

To establish the credibility of my research design, it is important that the evolution of outcomes in the control states approximate the counterfactual evolution of outcomes in the treatment states absent RTW legislation: more specifically, pre-intervention trends in our variables of interest have to be more or less parallel for the two groups of states. Figure 3.1 juxtaposes the raw trends in the control states against trends in each of the treatment states for a variable of the utmost interest, union coverage rate ('coverage'). Figures comparing raw trends in other variables are relegated to Appendix B.2.

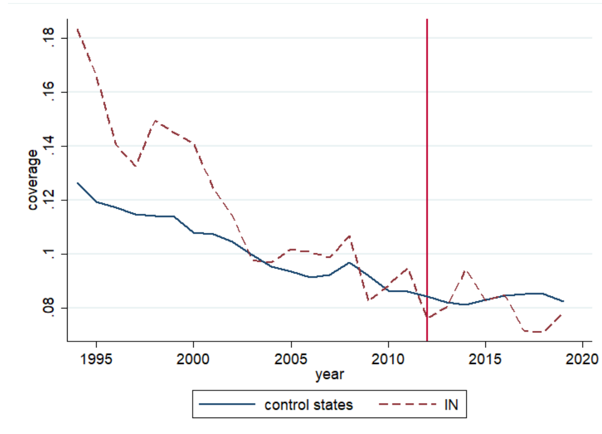
It can be seen from Figure 3.1 that trends in union coverage are not exactly parallel over the long run: in IN, MI, and WI, coverage rates appear to have declined somewhat more steeply than in the control states since the 1990's. This is something almost to be expected in long panels: small differences in trends that are not detectable in short panels may become more visible when they build up over sufficiently long periods. This can be a concern for causal identification if we are trying to estimate long-term effects. But given that the length of exposure to RTW laws for some of the treated states in my sample is as short as three years, and also given that OK (which has the longest post-treatment period) mostly serves as a control unit once other states become exposed to treatment, the RTW coefficient in (3.1) will be identified mostly by short-term effects of fewer than eight years (see Goodman-Bacon, 2018 for the details of this argument). To gauge such short-term effects, it would suffice to ensure that trends are more or less parallel over a period immediately preceding exposure to treatment. To do this, I estimate the following event-study equation:

$$y_{st} = a_s + b_t + \sum_{\tau=-23}^{18} \beta_\tau D_{s,\tau} + \epsilon_{st}, \tag{3.2}$$

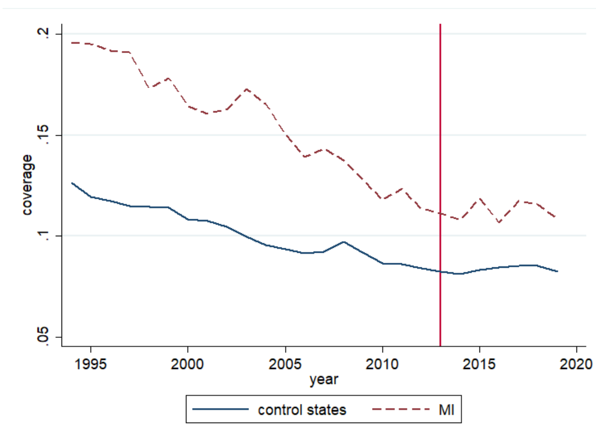
where $D_{s,\tau}$ is a dummy indicating that state $s$ in year $t$ is $\tau$ years from its initial year of exposure to a RTW law ($D_{s,\tau}$ for $\tau = -1$ is omitted in the actual regression). The event study coefficients $\beta_\tau$ will thus capture the average 'effect' of being a treatment state $|\tau|$ years before/since RTW adoption; the slope of the series of $\beta_\tau$ in the pre-treatment years would pick up any differential trends exhibited by the treatment states as a whole relative to the control states. Plots of event study coefficients (along with 95 percent confidence



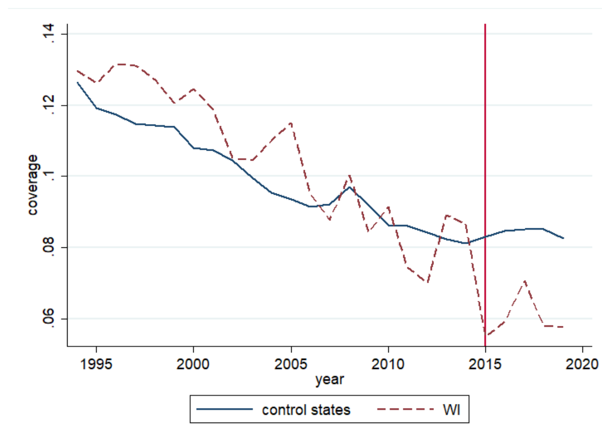

(a) Oklahoma




(b) Indiana

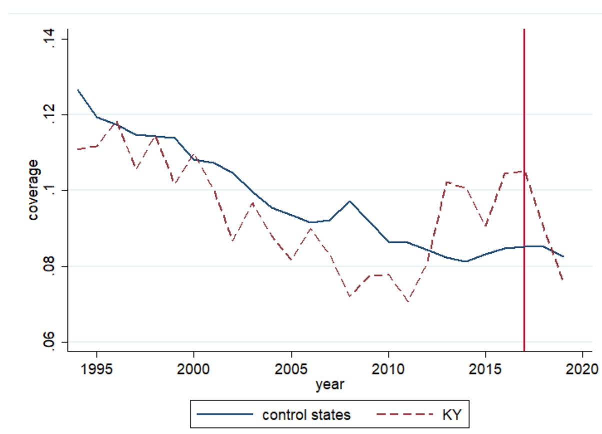



(c) Michigan




(d) Wisconsin


(e) West Virginia



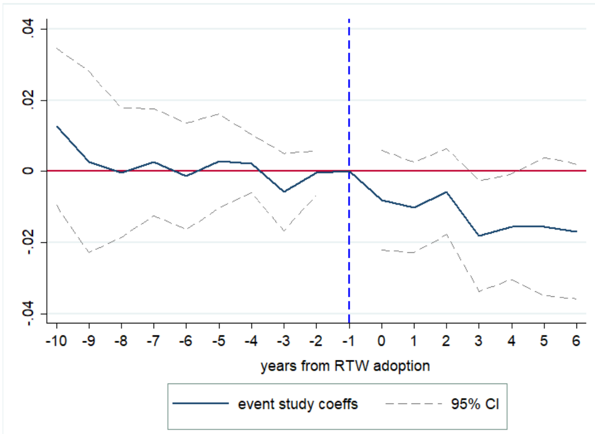

(f) Kentucky

Figure 3.1: Raw trends in union coverage

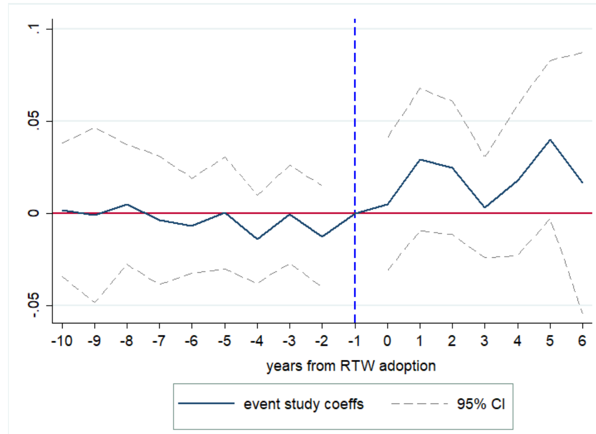
intervals) estimated using different left-hand-side variables in (3.2) are presented in Figure 3.2 (to economize on space, I omit plots for variables that are closely related to others, such as Dgdp and those pertaining to legislative control). The $\hat{\beta}_\tau$'s are plotted only for an event window spanning from $\tau = -10$ to 6, because I consider ten years running up to RTW enactment to be sufficient to detect locally divergent pre-trends, and because panels become highly unbalanced after $\tau = 6$: up to that point, event-study coefficients are identified by at least three states; beyond $\tau = 7$, the $\beta_\tau$ coefficients are only identified by OK.

Figure 3.2 reveals a number of interesting patterns. First, it is evident that some variables trended differently in the treatment states compared to the control states: non-union wages, all macroeconomic variables, and all variables pertaining to political climate. Economic activity as measured by GDP per capita, employment rate, and share of manufacturing employment all appear to have slowed in the treatment states relative to the control states prior to RTW legislation, which may also explain the relative fall in non-union wages in those states. At the same time, measures of liberal-leaning (in the American sense of the term) political climate also relatively declined in the treatment states. These patterns suggest that declining support for Democrats, coupled with economic hardship, facilitates passage of RTW laws.

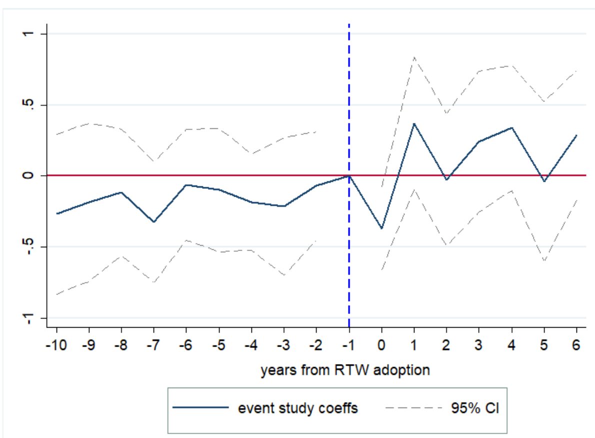
Secondly, all economic performance measures (other than wages) appear to improve in the post-treatment period, seemingly vindicating the claim of RTW proponents that these laws save jobs. Should we then discard the macroeconomic variables as improper controls? (or better yet, focus on studying the macroeconomic impacts of RTW laws?) I will argue otherwise. GDP and employment are known to behave cyclically, and it is easy for laws passed near the trough of a cycle to be spuriously correlated with the subsequent economic recovery. At least in the case of IN and MI, the pre-RTW decline and then rebound in relative economic performance may be better explained by the heavy concentration of manufacturing in those states – which presumably made them more vulnerable to the shock of the



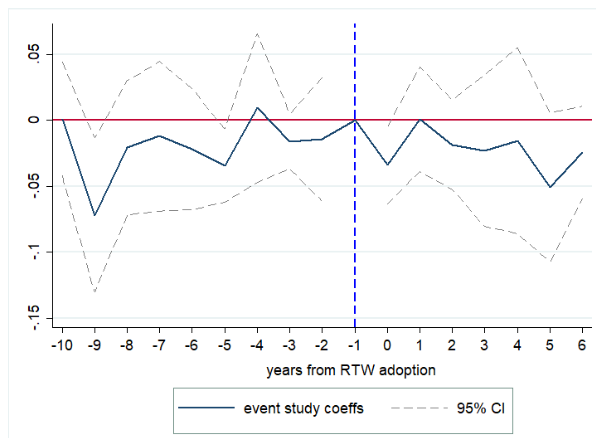

(a) dep var: coverage



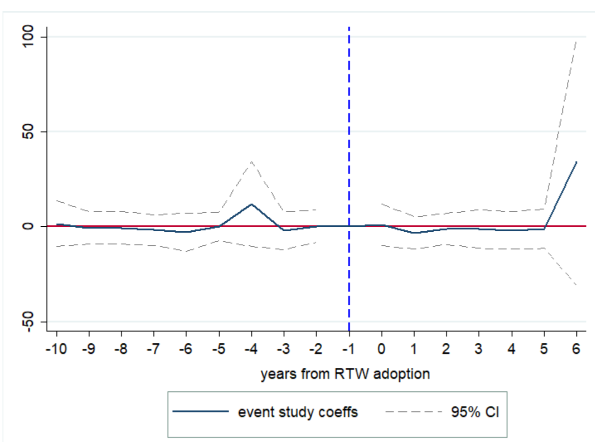

(b) dep var: freeride



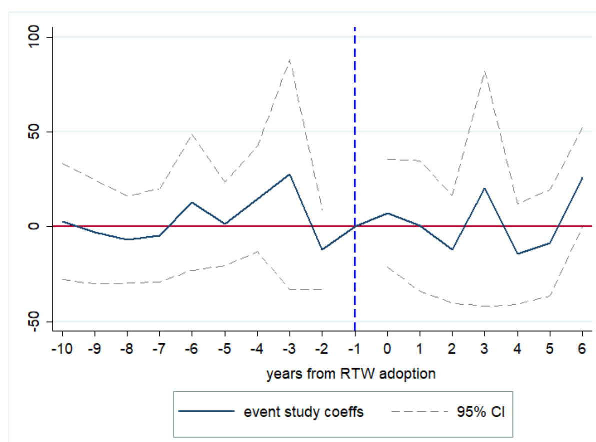

(c) dep var: inflow_perK




(d) dep var: outflow_perK

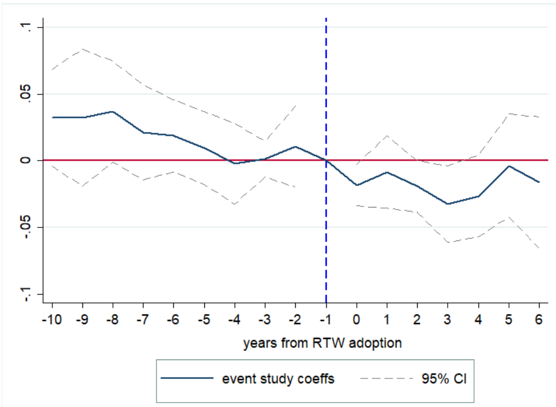



(e) dep var: strikehazard

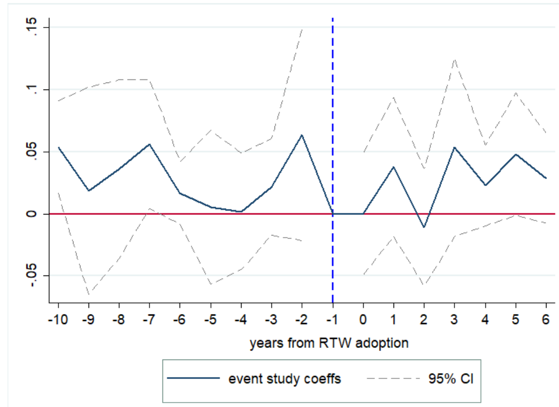



(f) dep var: duration

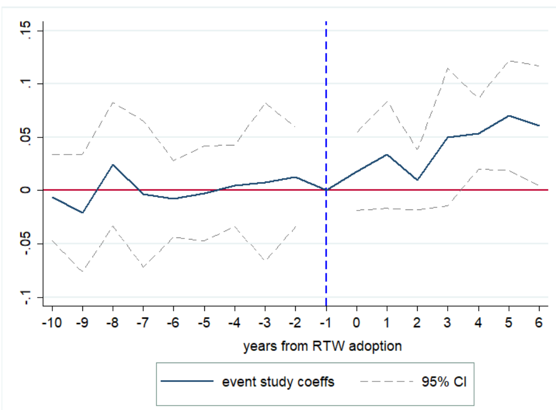
Figure 3.2: Event studies of RTW effects



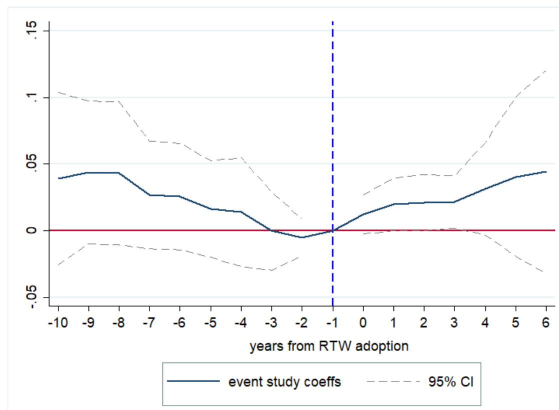

(g) dep var: log of nuwage



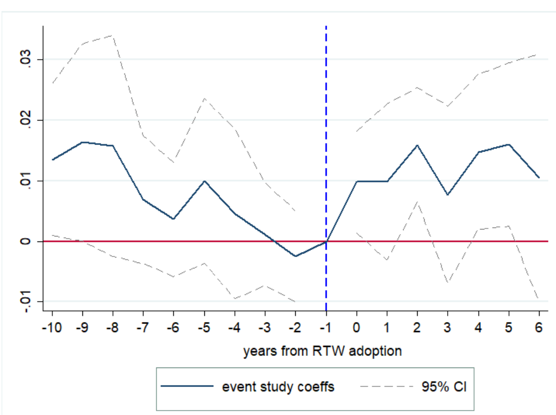

(h) dep var: log of union wage



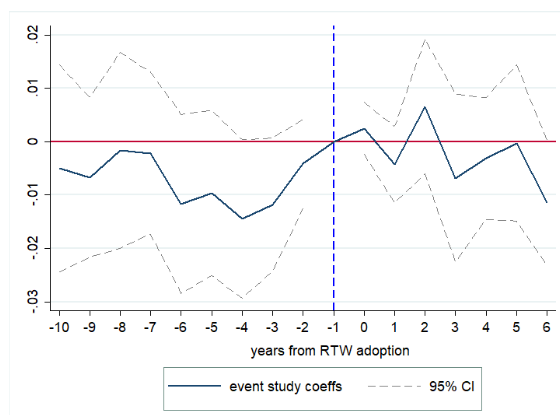

(i) dep var: udiff




(j) dep var: log of gdp_percap

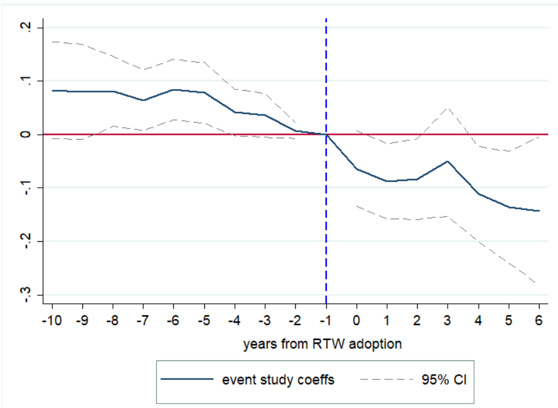



(k) dep var: emp

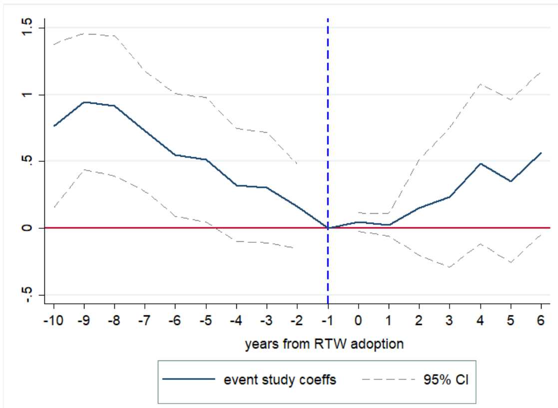



(l) dep var: manuf

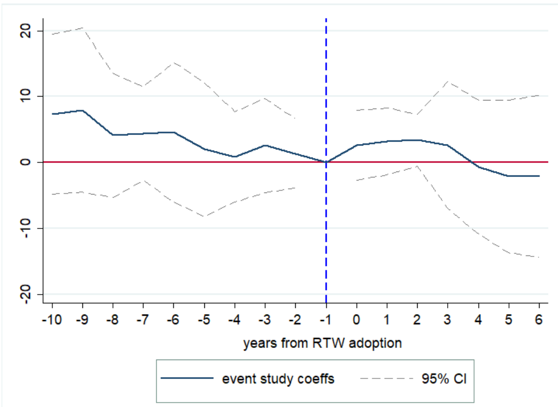
Figure 3.2: Event studies of RTW effects (continued)



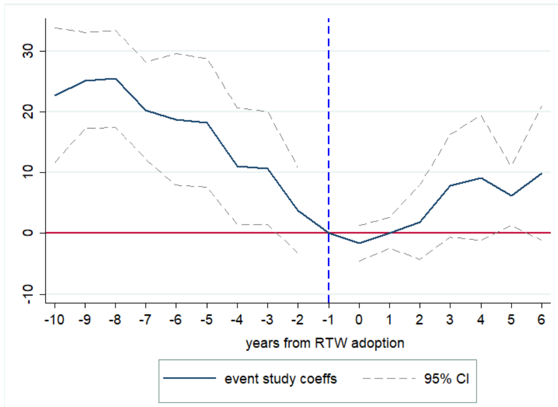

(m) dep var: house_dem




(n) dep var: dgov




(o) dep var: ctz_ideo




(p) dep var: gvt_ideo

Figure 3.2: Event studies of RTW effects (continued)

Great Recession but also enabled them to recover faster once global demand for their goods eventually picked up. Even if economic performance *were* affected by RTW laws, there is good reason to include them as controls. Economic conditions are very likely to affect union behavior; if so, RTW coefficients will in part capture the indirect effects RTW laws have on union behavior through their effects on employment, etc. If we are interested in the direct effects of RTW laws on union behavior instead of such reduced-form effects, it makes sense to net out the indirect effects by controlling for economic conditions.

Inclusion of the political variables may be justified on the same grounds: if RTW laws do help to entrench Republican rule, and if the Republicans, thus empowered, do things to undermine organized labor that they would have done with or without RTW laws, RTW coefficients will pick up the usual effects of Republican rule which may not really interest us. A remaining concern is whether conditioning on these variables would permit an apples-to-apples comparison. States adopting RTW laws may be different from non-RTW states with similar observed political-economic trends: perhaps the ideology of Republican leaders in such states were radicalizing in unobserved ways, which may have driven the passage of RTW laws. My hope is that the measures of citizen and government ideology would be a good enough proxy for such developments to address this concern.

Returning once again to Figure 3.2, we find that event study coefficients for the outcome variables freeride, inflow_perK, and udiff visibly increase in the post-treatment period, but also appear to trend slightly upward in the pre-treatment period (the opposite pattern holds for coverage). Even if these suspected differential pre-trends do not prove statistically significant, if they exist they may bias upward (downward) RTW effect estimates based on the baseline DID model in (3.1). Thus it would be prudent to additionally control for time trends (either group-specific or state-specific) as a check for robustness.

As a parametric test of the parallel-trends assumption, I estimate the following equation:

$$y_{st} = a_s + b_t + \gamma\tau_{st} + \epsilon_{st}, \quad \tau_{st} < 1. \tag{3.3}$$

where $\tau_{st}$ is zero for all control states; for a treatment state, it is the number of years from initial exposure $t$ (e.g. for OK, $\tau = \{-7, -6, ..., 0\}$). Estimates of $\gamma$ for select variables are presented in Table 3.2. As anticipated by the event study plots, we see that wages and the power of the Democratic Party in the treatment states trended more negatively than in the control states. Although not statistically significant, lgdp also appears to have moved in lockstep with wages. The significantly negative differential trend in inflow_perK indicates that unionization has become more difficult at a faster rate in the treatment states, perhaps because their depressed economies (and the higher threat of unemployment) made it more challenging for unions to overcome employer resistance when attempting to organize new workplaces.

Table 3.2: Parallel trends test results

| | coverage | freeride | inflow_perK | outflow_perK | udiff |
|---|---|---|---|---|---|
| $\tau$ | -0.001 | 0.000 | -0.099** | -0.002 | -0.000 |
| | (0.001) | (0.001) | (0.036) | (0.001) | (0.001) |
| | luwage | lnuwage | strikehazard | duration | lgdp |
| $\tau$ | -0.004** | -0.003** | -0.440 | 0.432 | -0.003 |
| | (0.002) | (0.001) | (0.377) | (0.538) | (0.002) |
| | emp | manuf | college | house_dem | gvt_ideo |
| $\tau$ | -0.001 | -0.000 | -0.001 | -0.009*** | -0.740** |
| | (0.001) | (0.000) | (0.000) | (0.002) | (0.274) |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

I now proceed to estimate the DID equation in (3.1) by progressively augmenting the vector of controls. I start with no controls other than the state- and year- fixed effects, then include the economic variables, then the political variables, and finally add controls for state-specific linear time trends. This last specification serves as the most stringent test of causality in my analysis: I will consider RTW effect estimates that prove robust to controlling for time trends as safe to interpret causally. The RTW coefficients obtained through this procedure for each outcome of interest are summarized in Table 3.3. Detailed regression tables for each outcome variable that include coefficients on the control variables are in Appendix B.2.

Table 3.3: RTW coefficients by outcome variable and model specification

| | controls added | | | |
|---|---|---|---|---|
| | none | economic | +political | +time trends |
| coverage | -0.014** | -0.016*** | -0.015*** | -0.009** |
| | (0.007) | (0.006) | (0.005) | (0.004) |
| lcoverage | -0.164*** | -0.178*** | -0.163** | -0.150*** |
| | (0.051) | (0.054) | (0.059) | (0.050) |
| freeride | 0.030 | 0.026 | 0.017 | 0.015 |
| | (0.018) | (0.017) | (0.013) | (0.015) |
| inflow_perK | -0.049 | 0.049 | 0.288 | 0.282* |
| | (0.270) | (0.232) | (0.209) | (0.166) |
| outflow_perK | -0.012 | -0.010 | -0.025 | -0.036** |
| | (0.008) | (0.009) | (0.020) | (0.016) |
| udiff | 0.020 | 0.022* | 0.021* | 0.036*** |
| | (0.017) | (0.012) | (0.012) | (0.012) |
| luwage | -0.025 | -0.017 | -0.008 | 0.032 |
| | (0.020) | (0.021) | (0.024) | (0.024) |
| lnuwage | -0.034** | -0.024** | -0.024** | -0.004 |
| | (0.014) | (0.010) | (0.010) | (0.011) |
| wdiff | 0.009 | 0.008 | 0.016 | 0.036 |
| | (0.016) | (0.016) | (0.019) | (0.024) |
| strikehazard | -0.157 | 2.656 | 4.977 | 3.286 |
| | (3.698) | (3.230) | (4.656) | (4.690) |
| duration | -3.477 | -3.784 | -1.680 | -2.533 |
| | (4.971) | (5.304) | (7.679) | (10.557) |

Robust standard errors in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

The result that most stands out from Table 3.3 is that RTW laws *do* seem to reduce the extent of unionization. Estimates of RTW effects on union coverage are significant and robust across specifications; even after controlling for time trends, union coverage declines by nearly one percentage point (from a baseline of 10 percent), or by 15 percent when coverage is measured in logs, under a RTW law. In percentage points, this is fairly close to Eren and Ozbeklik (2016)'s estimated effect size of 1.3 percentage points in OKlahoma's case (although their outcome measure is union density, not coverage).

An equally striking set of results is that under RTW laws, (i) free-riding increases very little, if at all (at most three out of a hundred covered workers forego union membership); (ii) inflows of newly unionized workers, rather than decreasing, may in fact increase; and (iii) outflows through de-certification elections, if anything, *decreases*. Neither the Free-rider Hypothesis nor the Bargaining Power Hypothesis predict such results: declining union bargaining power under RTW ought in theory to lead to more frequent union de-certifications if workers become dissatisfied with their union's performance (or if employers sense greater vulnerability in unions). The positive estimates of effects on inflows contrast sharply with Ellwood and Fine (1987)'s finding, as well as contradicting the predictions of a decline in either the supply or demand for unionization under the Free-rider Hypothesis.

The natural question that arises is: what explains these unexpected results, and how can it be that RTW laws reduce union coverage without inducing either an increase in free-riding or a decline in unionization activity?

The insignificant RTW effect on freeride could mask heterogeneous effects among treatment states: looking at the raw trends in this variable displayed in Appendix B.2, it can be seen that nonmember shares increased quite noticeably in OK and KY in their post-treatment periods. Even in these cases, however, the increase amounts to little more than 10 percentage points. Such a weak response in free-riding may be explained in a number of ways: First, most workers who join unions may value membership by more than the fees associated with it, so that the opportunity to save on the fees entices only a few uncommitted members to leave; second, unions may respond to RTW laws by heightened efforts to recruit members and instill loyalty; third, complementing the previous explanations, the force of social custom and solidarity may be a much bigger determinant of union membership than pecuniary considerations (in that regard, *homo economicus* may be a particularly ill-suited model for analyzing the behavior of unionists).

An alternative interpretation could be that the true extent of RTW-induced free-riding is not fully captured by changes in the share of covered nonmembers. Strictly speaking, a free-rider is only observed conditional on there being a union representing the individual. If a union ceased to exist as a result of rampant free-riding, we are unable to observe ex post the free-riders who were responsible. Moreover, if unions that dissolve in this manner had high shares of nonmembers to begin with, their exit could even pull down the average share of nonmembers among the surviving unions. Thus the RTW effect estimator may be biased downward: under the extreme assumption that all 15 percent of unions 'killed' by RTW had nonmember shares close to 100 percent, this bias could be as large as 0.15 (15 out of 100 covered employees); a more reasonable magnitude could be 0.07, so that an estimated effect of 0.03 would indicate a true effect size of 0.1. While additional free-riding by ten percent of covered employees is far from trivial, it would not by itself amount to a death sentence for an average union that signs up 90 percent of covered employees. But a union with a sign-up rate of only 50 percent may face a more existential threat: in addition to membership dropping to 40 percent, agency fee revenue from the remaining 60 percent of employees will suddenly be lost. If it is mainly these types of unions that dissolve under RTW laws, then RTW effects on union coverage can be thought to operate mainly through removal of an existing source of revenue than through distortion of incentives. In other words, it would be the existing stock of nonmembers, not newly added nonmembers, that causes such unions' downfall. The question might then be asked, why were these unions afflicted by high levels of non-membership even before RTW legislation? If the answer is that they did not deliver much for their workers and so failed to inspire loyalty, their disappearance may not be very concerning; it would be more concerning if the answer has to do with objective obstacles to recruiting new members (e.g. geographic dispersion of work sites). Without being able to observe outcomes at the level of individual unions, however, all of this is conjectural and there is little more we can say about the mechanisms of union attrition under RTW laws.

The 'wrong' signs of RTW coefficients for the flow variables (inflow_perK and outflow_perK) may be explained thus: when nonmembers cannot be forced to contribute financially to a union that represents them, it becomes harder for an employer to persuade employees against voting for union representation by claiming that a union would steal from their paychecks. At the same time, already-unionized workers would lose an incentive to try to vote out a union they dislike, because they now have the option to simply leave the union and stop contributing to a public good which they feel is overpriced. This is a scenario where incentives to free-ride may paradoxically work to the benefit of unions: workers who perceive an opportunity to free-ride on union services may be both more willing to vote for unionization and more reluctant to vote out an existing one.

Whether RTW laws actually *increased* the inflow of unionized workers is uncertain. The estimated effect sizes are not stable across specifications and only become marginally significant when all controls are included. But given that inflow_perK had trended negatively in the treatment states (at a rate of almost 0.1 workers per thousand: see Table 3.2), it is not surprising that controlling for observables that may explain such a trend would alter the coefficients. Once political variables are accounted for, the effect size is invariant to controlling for time trends – the main effect of including time trends appears to be to increase the precision of the estimate. And the effect size is quite large: from a baseline of 0.76 non-union workers per thousand[8] becoming unionized per year, an increase by 0.28 workers per thousand translates to a 37 percent increase. Also, the event study of effects on union formation in Figure 3.2c show a jump in the series of event-study coefficients immedi-

[8]this is the mean of inflow_perK for treatment states over a five-year period preceding RTW legislation

ately following year zero, which suggests a causal effect. Taken together, these results may be viewed as weak but suggestive evidence that RTW laws facilitate new union organizing. Although my tests are underpowered to firmly reach such a conclusion, perhaps a follow-up study using updated data on NLRB elections may be able to. The same cannot be said about the effect on outflows. The only significant RTW coefficient for this outcome seems to be an artefact of fitting a linear time trend to data points that are not-so-linear. The event study of RTW effects on de-certifications in Figure 3.2d reveal no sharp discontinuities or trend breaks around year zero to suggest a causal effect.

Hence RTW laws could not have caused a decline in union coverage by reducing the rate of union formation. Neither did they increase the rate of union destruction through the formal channel of de-certification elections. That leaves us with two possibilities: RTW laws could have either accelerated the closures of unionized establishments (in which case unions die a natural death) or driven unions to premature deaths outside the de-certification channel. The former possibility is unlikely given the avowed support of business interests for RTW legislation: it is doubtful that corporate lobbies such as Americans for Prosperity would back a policy that can lead to widespread plant closures. The second possibility is one that was previously alluded to: unions that were already in a precarious position owing to high nonmember shares could have been disproportionately affected by the loss of agency fee revenue. Such unions may have effectively ceased to function as bargaining agents, and let existing contracts expire without negotiating new ones.

As anticipated by the event study plots, RTW laws are found to increase union wage differentials by as much as 3.6 percentage points. In specifications that do not control for state-specific time trends, the coefficients are smaller and only significant at the ten percent level. But these are likely to be underestimates: Goodman-Bacon (2018) demonstrates that when a treatment effect is a trend-break rather than a level shift, a two-way fixed-effects DID estimator is biased toward zero, and we see such a trend-break in the event study of RTW effects on udiff (Figure 3.2i). Therefore I consider estimates of a 2.1 to 2.2 percentage point increase in udiff as a lower bound. Moreover, these summary measures can mask effects that are concentrated in time: between the fifth and seventh years of exposure to a RTW, the effect can be larger than 5 percentage points.

Again, this is a result that contradicts conventional hypotheses. Precisely how RTW laws can cause this phenomenon is unclear. As mentioned in Section 3.3, one possibility is that insecure unions try to increase their appeal among workers by extracting greater concessions from employers (this may be termed a behavioral effect). Another is that the attrition of weaker unions with low union wage differentials pulls up the average differential of the surviving unions (a compositional or selection effect, similar to one hypothesized by Farber, 1984). This channel is plausible if unions with low wage premia also tend to have high shares of nonmembers, which, as I conjectured, makes them particularly vulnerable to RTW laws. Schumacher (1999) in fact finds evidence that nonmember shares are higher in sectors with low union wage differentials.

I judge it unlikely, however, that the increase in udiff is driven purely by compositional change. In order for the selection effect alone to account for a 5 percentage point increase in udiff (from a baseline of 20 percent), while union coverage falls by 2 percentage points (from a baseline of 10 percent), the unions that fall victim to RTW must have had an average udiff close to zero – a rather implausible assumption. It would be more plausible to suppose there is at least *some* behavioral component to the response in the union wage differential: agents do respond to incentives, and unions faced with the threat of membership loss would be acting rationally by trying to prove their worth to their constituents, including by winning higher pay raises. This would help to make sense of Makridis (2019)'s finding of increased satisfaction on the job for union members after exposure to RTW laws. Perhaps such a behavioral effect also helps to explain our tentative evidence of a rise in the rate of union formation under RTW: unions in the U.S. have for a long time engaged in 'concession bargaining' in which gains won in the past were continually given up in the name of saving jobs. Concessions made to employers, such as the infamous two-tier wage system in the Big Three U.S. automakers,"sullie[d] the reputation of unions as bargaining agents, making bargaining seem like a losing deal in the eyes of non-union workers" (Chaison, 2012). If RTW laws jolted unions into reversing course and taking back lost ground, non-union workers may well have rediscovered unionism as a solution to their problems.

Turning to the RTW coefficients for union and non-union wages, first note that estimated effects on lnuwage are consistently more negative than those on luwage, which dovetails with the finding of a positive effect on udiff (the difference between the RTW coefficients for luwage and lnuwage is equivalent to RTW coefficients for the unconditional union wage gap, which for illustration is generated as a new variable 'wdiff'). Both sets of coefficients become more positive with the addition of controls, especially with the inclusion of time trends: the coefficient for luwage turns positive as a result, and the coefficient for lnuwage loses significance. This pattern suggests there exists a persistently negative differential trend in wages in the treatment states that cannot be explained by observed economic or political developments, and that what appears to be an RTW effect is just a continuation of that same trend. I cannot think of an obvious omitted variable that may account for this trend. Increasing exposure to competition from low-wage countries may be one, but to the extent this affects wages and corporate earnings in an equal manner, the effect should be captured by lgdp. Another candidate is growing regional disparities in costs of living, which may generate differential pressures on wages. Lastly, repeated rounds of 'concession bargaining' by major unions rooted in the treatment states may have eroded both union and non-union wages (by degrading wage patterns) at a faster rate than in the rest of the country.

The last two rows of Table 3.3 indicate lack of evidence that RTW laws affect strike activity. All RTW coefficients for measures of strike activity are statistically zero. If unions did become more assertive in their demands under RTW, these results suggest that in most cases they were able to win their demands without walking off the job. A possible implication is that there may have been considerable room for negotiation around wages that unions were not sufficiently exploiting prior to RTW.

Thus far, we have two results that support the existence of causal effects of RTW laws: a negative effect on union coverage and a positive effect on union wage differentials. These results are fairly robust to removing any single treatment state from the sample. Table 3.4 presents the RTW coefficients obtained from DID regressions controlling for all available covariates (including time trends), each time removing one treatment state from the sample.

Table 3.4: Sensitivity of RTW coefficients to leaving out one RTW state

| | treatment state removed from sample | | | | | |
|---|---|---|---|---|---|---|
| | OK | IN | MI | WI | WV | KY |
| coverage | -0.010** | -0.011** | -0.009* | -0.005* | -0.007* | -0.009** |
| | (0.005) | (0.004) | (0.005) | (0.003) | (0.004) | (0.004) |
| lcoverage | -0.156** | -0.171*** | -0.154** | -0.101*** | -0.141** | -0.168*** |
| | (0.063) | (0.054) | (0.061) | (0.028) | (0.056) | (0.054) |
| udiff | 0.040*** | 0.039*** | 0.037** | 0.029** | 0.039*** | 0.028** |
| | (0.013) | (0.013) | (0.015) | (0.013) | (0.013) | (0.012) |

Robust standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

It can be seen from the first row that Wisconsin is an influential unit: without it, the estimated RTW effect on union coverage shrinks to nearly half of the -0.01 percentage point effect estimating using all treatment states. In percent terms as well, the RTW coefficient drops from 15 percent to 10. With regard to RTW effect on udiff, WI and KY each appear to be influential. But it is also seen that the significance of our baseline results do not depend solely on these outlying states.

As a further test of robustness, I re-estimate RTW effects on coverage and udiff using an alternative approach. The synthetic control method (SCM) due to Abadie et al. (2010) implements fundamentally the same idea as a DID, namely to compare the evolution of an outcome in a treated unit to its counterfactual evolution in the absence of treatment, where this counterfactual is approximated by outcomes in untreated units. The difference is that whereas DID compares the evolution of the treated unit to that of each available control unit, SCM synthesizes a single counterfactual unit (the synthetic control) from the pool of available units based on how close they are in their characteristics and paths of evolution to the treated unit – assigning greater weights to more similar units and zero weights to those that are too different. This feature allows SCM to overcome two limitations of the DID approach, namely its restrictive requirement of parallel trends between the treated and untreated units, and potential bias arising from linear extrapolation between drastically different units. A weakness of SCM is that it does not produce reliable estimates when a synthetic control built from the available pool of control units does not match closely enough the pre-treatment outcomes and characteristics of the treated unit(s).

I obtain SCM estimates of RTW effects on coverage and udiff using the Stata synth_runner package written by Galiani and Quistorff (2017), which implements an extension by Cavallo et al. (2013) that allows for multiple units to experience treatment. The SCM estimator in this case averages the outcome differences between each pair of the treated unit and its synthetic control, for each post-treatment period. The associated p-values are derived based on permutations that assign placebo treatments to each available control unit. To examine the trajectory of RTW effects over more than three years, I use only four treatment states (excluding WV and KY) that have longer post-treatment periods. I employ two different ways to match control units to the treated: matching on levels of pre-treament outcomes and matching on pre-treatment trends. When matched on trends, the resulting SCM estimate can be interpreted as a percent change. In either case, the predictor variables used for matching were automatically selected so that the donor states assigned to the synthetic control on the basis of their values of those predictor variables over a 'training' phase (covering the initial 80 percent of the pre-treatment period) best predict the outcome path of the treatment state during the 'validation' phase (the remaining 20 percent of the pre-treatment period).

Table 3.5: Synthetic Control Method estimates of RTW effects and p-values

| | coverage | | udiff | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| year 0 | -0.011 | -0.049 | 0.018 | 0.059 |
| | (0.115) | (0.614) | (0.461) | (0.467) |
| year 1 | -0.008 | -0.037 | 0.026 | 0.156 |
| | (0.297) | (0.714) | (0.213) | (0.118) |
| year 2 | -0.005 | 0.003 | 0.015 | 0.118 |
| | (0.285) | (0.947) | (0.402) | (0.369) |
| year 3 | -0.013 | -0.143 | 0.030 | 0.187 |
| | (0.076) | (0.214) | (0.081) | (0.107) |
| year 4 | -0.009 | -0.099 | 0.050 | 0.289 |
| | (0.210) | (0.272) | (0.243) | (0.209) |
| Trends matched | No | Yes | No | Yes |

Standardized p values in parentheses

Table 3.6: Event study coefficients of RTW effects

| | coverage | | udiff | |
|---|---|---|---|---|
| | level | log | level | log |
| year 0 | -0.009 | -0.095 | 0.020 | 0.101 |
| | (0.011) | (0.159) | (0.022) | (0.116) |
| year 1 | -0.008 | -0.083 | 0.020 | 0.108 |
| | (0.008) | (0.127) | (0.037) | (0.187) |
| year 2 | -0.003 | -0.041 | 0.009 | 0.080 |
| | (0.007) | (0.073) | (0.009) | (0.061) |
| year 3 | -0.015* | -0.214* | 0.042 | 0.226 |
| | (0.007) | (0.090) | (0.035) | (0.185) |
| year 4 | -0.010 | -0.184* | 0.043 | 0.270 |
| | (0.006) | (0.072) | (0.022) | (0.135) |
| $N$ | 675 | 675 | 675 | 672 |
| Adjusted $R^2$ | 0.666 | 0.630 | 0.125 | 0.087 |

Robust standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.5 presents the results. It can be seen that estimated effects on coverage and udiff have consistent signs across all years and specifications, and are of comparable magnitudes to their event-study counterparts. Table 3.6 presents the corresponding event study coefficients from a regression using the same four states as the treated units, and controlling for all available economic and political variables. The similarities of estimates across the two methods suggest that any extrapolation bias in the DID is not likely to be large.

Based on the weight of the evidence thus far, I conclude that the RTW laws in the six recent-adopting states likely had a negative effect on union coverage and a positive effect on union wage differentials. It is unlikely that RTW laws exerted such an effect on coverage by inducing existing union members to leave, given the finding of a null effect on free-riding.

## 3.5 Public Sector Results

In this section, I perform much the same analyses as in the previous section using the public sector sample. One thing to anticipate is that any RTW effect is likely to be confounded by other laws that went into effect in some of the treatment states prior to RTW adoption, which placed severe limits on the rights of public sector unions. For example, in March 2011 Gov. Scott Walker of Wisconsin signed into law Act 10 which prohibited state- and municipal-government employee unions from bargaining over hours, working conditions, and compensation other than base pay; mandated annual union re-certification elections; and prohibited union dues deductions from paychecks (which is one type of a union security arrangement). Michigan in March 2011 enacted Act No. 9 giving local government-appointed 'emergency managers' the power to nullify public sector union contracts. Indiana abolished state employees' collective bargaining rights by executive order in 2005. In a study of four major public sector unions in Wisconsin, Nack et al. (2020) find that Act 10 led to a 70 percent reduction in active membership; media reports suggest that the impact of Indiana

Gov. Mitch Daniels' 2005 executive order on union membership was equally devastating.[9] The Wisconsin and Michigan laws were part of a nationwide wave of legislative attacks on public sector unionism in the years 2011-12, which saw 15 state legislatures passing similar laws (Lafer, 2013). The scope and detail of these laws vary, and I do not attempt to identify their separate impacts.

Another peculiarity of industrial relations in the U.S. public sector is that strikes are illegal for teachers and public safety workers in 39 states, while 12 states allow strikes for most categories of workers except police and firefighters (Sanes and Schmitt, 2014). This makes public sector strikes a rare event: the mean annual number of public sector strikes in a state is 0.25, and out of a total of 343 strikes observed over the 26-year sample period, a handful of states account for the vast majority (CA, 55; IL, 122; OH, 57). This makes statistical inference inherently challenging. It also means that my strike duration variable is undefined and treated as missing in numerous cells (if no strike is observed in a given year and state, it is hard to say how long one would have lasted). I will therefore exclude strike duration from my outcome variables of interest and focus on the remaining seven.

One final complication in analyzing the effect of RTW laws in the public sector has to do with the fact that the U.S. Supreme Court decision in the *Janus v. AFSCME* case has made the entire U.S. public sector effectively RTW as of July 2018. This is potentially a different kind of treatment than RTW laws legislated at the state level, the effects of which warrant a separate examination. Hence I attempt to measure separately (as well as together) the effects of RTW on public sector unionism in the original six treatment states and in the 24 non-RTW states (including DC) that become RTW (as far as the public sector is concerned) in 2018. This necessitates a somewhat confusing redefinition of treatment and control states for the purpose of estimating the impacts of *Janus v. AFSCME* only: the

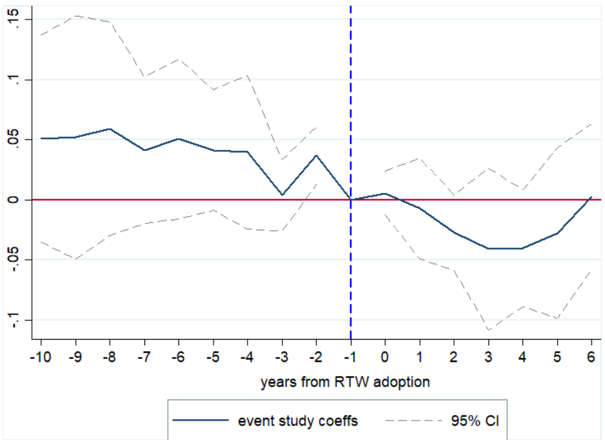
[9] Unnamed author (2011, March 10) *In Indiana, a case study in ending collective bargaining.* Wisconsin State Journal `https://madison.com/wsj/news/local/govt-and-politics/in-indiana-a-case-study-in-ending-collective-bargaining/article_badb25e6-4b71-11e0-97c4-001cc4c002e0.html` (accessed September 2020).

treatment states in this case are the 24 'Never RTW' states that served as the control group in the private sector analysis, while the 'Always RTW' states now comprise the control group. There are both advantages and drawbacks to a separate accounting of *Janus v. AFSCME*'s effect. An advantage is that for the 24 states that were affected, the regime switch was plausibly exogenous and therefore RTW effects are less likely to be confounded by political developments specific to those states. A disadvantage is that the treated states had less than two years of exposure to the new regime before the 2020 pandemic: if RTW effects grow over time, estimates based on those two years would understate the true effects over a policy-relevant time frame.

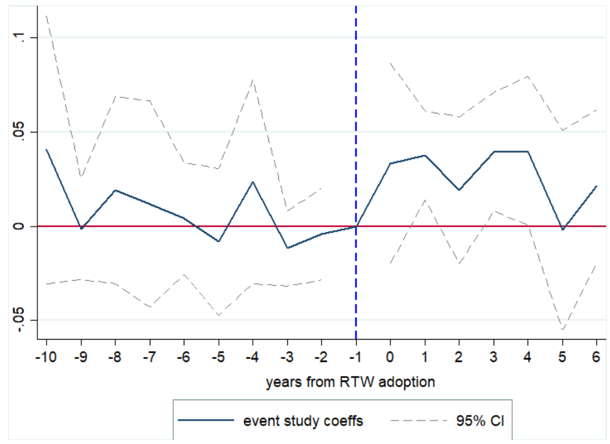
I present summary statistics for six variables of interest in Table 3.7. Figure 3.3 presents event studies of RTW effects on these variables. As the point here is to see whether the states that voluntarily adopted RTW laws were different with regard to levels or pre-trends in the outcome variables, I limit the treatment group to the original six treatment states. I do not trim the sample period to pre-2018 years in the event studies, in order to maximize panel balance and account for possible time-varying treatment effects.

From Table 3.7 we learn that public sector unions in the original treatment states were in an embattled position: unionization rates are closer to Always RTW states than to other non-RTW states, nonmember shares are higher, nominal wages are much lower, and union wage differentials are halfway between Always RTW and Never RTW states. Strike participation levels appear much higher among the switcher states: this is driven by the incidence of the 2017-2018 wave of teacher strikes which affected WV, KY and OK, among others; prior to 2017, average strike participation in the treatment states was closer to one in a thousand unionized workers.

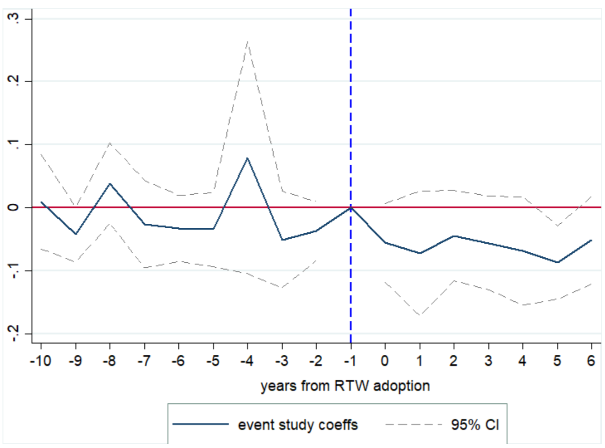
Several things can be learned from Figure 3.3. First, panel (a) shows that public sector union coverage declined (relative to control states) in the treatment states after RTW legislation, but that a negative pre-trend may have preceded RTW laws by a few years: the event-study



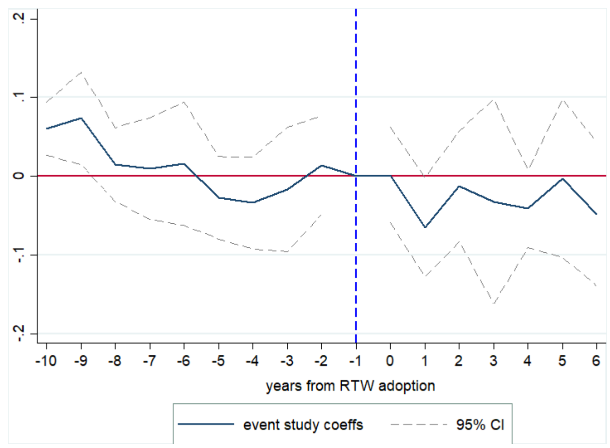

(a) dep var: coverage



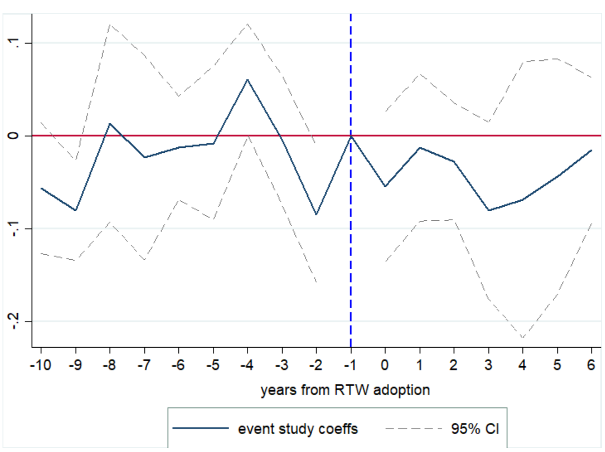

(b) dep var: freeride



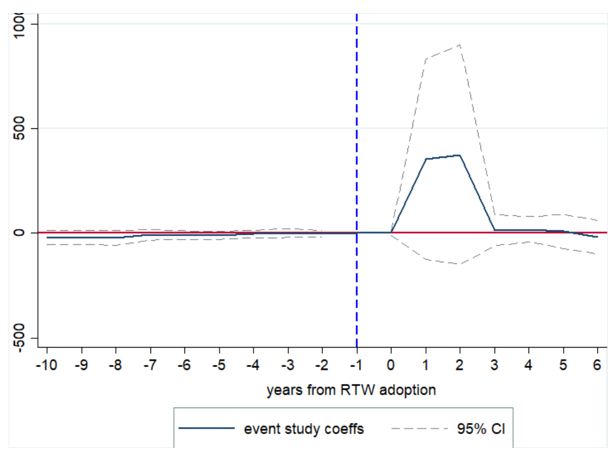

(c) dep var: log of uwage




(d) dep var: log of nuwage




(e) dep var: udiff




(f) dep var: strikehazard

Figure 3.3: Event studies of RTW effects, public sector

Table 3.7: Summary statistics, public sector

| | Never RTW | Switchers | Always RTW |
|---|---|---|---|
| coverage | 0.526 | 0.340 | 0.216 |
| | (0.142) | (0.144) | (0.095) |
| freeride | 0.087 | 0.133 | 0.219 |
| | (0.067) | (0.077) | (0.091) |
| uwage | 22.241 | 19.674 | 19.360 |
| | (5.314) | (4.016) | (4.477) |
| nuwage | 19.176 | 16.410 | 16.611 |
| | (4.935) | (3.421) | (3.821) |
| udiff | 0.117 | 0.104 | 0.085 |
| | (0.099) | (0.082) | (0.092) |
| strikehazard | 6.178 | 36.520 | 11.711 |
| | (38.407) | (215.564) | (153.223) |
| duration | 16.208 | 17.747 | 4.900 |
| | (21.768) | (21.825) | (4.581) |

Means and standard deviations (SD in parentheses)

coefficients were already 0.05 points down from previous years in $\tau = -3$ and $\tau = -1$. Panel (e) indicates that udiff took a noticeable plunge 2-4 years before the arrival of RTW laws. These patterns likely convey the effects of the attack on public sector unionism waged by some of the treatment states prior to RTW legislation. This underlines the challenge of identifying the causal effects of state-legislated RTW laws in the public sector – they may be indistinguishable from the effects of other similarly-timed policy interventions. Panel (b) exhibits movements in free-riding behavior that are more consistent with a causal effect: event-study coefficients jump immediately after treatment exposure and are even significant in some years. Panel (f) may appear to suggest RTW laws foment strikes, but the spike in years 1 and 2 is driven by WV and KY, which were part of the wave of teacher strikes in 2017 and 2018. It would be hard to make a connection between RTW laws and strikes based on just these two cases, unless there is reason to believe those strikes were directly motivated by lack of union security.

I now repeat the same DID regressions from Section 3.4 using three different samples: a trimmed sample containing observations up to 2017 and excluding Always-RTW states, a full sample including all 50 states (plus DC) up to 2019, and a narrowed-down sample that excludes the six original treatment states. RTW coefficients obtained from these samples would respectively capture effects on the original treatment states, all states that were non-RTW as of 2000, and the states that are made RTW by *Janus v. AFSCME*. Estimates from the trimmed sample are reported in Table 3.8, Table 3.9 reports estimates using the full sample, and Table 3.10 using the narrowed-down sample.

Table 3.8: RTW coefficients ignoring impact of *Janus v. AFSCME*

| | controls added | | | |
|---|---|---|---|---|
| | none | economic | +political | +time trends |
| coverage | -0.057* | -0.058 | -0.045 | -0.038* |
| | (0.033) | (0.034) | (0.029) | (0.020) |
| lcoverage | -0.169* | -0.166* | -0.132 | -0.119* |
| | (0.086) | (0.091) | (0.082) | (0.068) |
| freeride | 0.021 | 0.028* | 0.022 | 0.030 |
| | (0.017) | (0.015) | (0.019) | (0.029) |
| udiff | -0.016 | -0.024 | -0.028 | -0.031 |
| | (0.021) | (0.023) | (0.020) | (0.029) |
| luwage | -0.046** | -0.047** | -0.044 | -0.008 |
| | (0.021) | (0.018) | (0.026) | (0.032) |
| lnuwage | -0.055*** | -0.041** | -0.027 | 0.002 |
| | (0.016) | (0.017) | (0.021) | (0.027) |
| strikehazard | 1.416 | 0.744 | -3.378 | -3.591 |
| | (1.794) | (1.916) | (3.967) | (4.614) |

Robust standard errors in parentheses.
*** p<0.01, ** p<0.05, * p<0.1
Note: regression sample excludes 'Always RTW' states and observations for 2018 and 2019.

Two patterns emerge from a comparison of the three tables. First, estimated effects on coverage and free-riding vary depending on whether the RTW law was legislated or imposed by the Supreme Court. In the former case, the effect on public sector union coverage and non-membership appears to be similar-sized to that in the private sector (a decline of more

Table 3.9: RTW coefficients reflecting impact of *Janus v. AFSCME*

| | controls added | | | |
|---|---|---|---|---|
| | none | economic | +political | +time trends |
| coverage | -0.048 | -0.049 | -0.041 | -0.021 |
| | (0.030) | (0.031) | (0.028) | (0.015) |
| lcoverage | -0.123 | -0.125 | -0.092 | -0.075 |
| | (0.083) | (0.085) | (0.082) | (0.056) |
| freeride | 0.020 | 0.024 | 0.021 | 0.007 |
| | (0.018) | (0.017) | (0.018) | (0.015) |
| udiff | -0.014 | -0.018 | -0.020 | -0.020 |
| | (0.018) | (0.019) | (0.019) | (0.014) |
| luwage | -0.041** | -0.044*** | -0.036*** | -0.015 |
| | (0.017) | (0.010) | (0.012) | (0.011) |
| lnuwage | -0.049*** | -0.040** | -0.035* | -0.010 |
| | (0.014) | (0.019) | (0.017) | (0.020) |
| strikehazard | -24.369 | -28.100 | -41.559 | -66.180 |
| | (20.663) | (22.340) | (34.555) | (58.322) |

Robust standard errors in parentheses.

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Note: regression sample includes ‘Always RTW’ states.

Table 3.10: RTW coefficients identified exclusively by *Janus v. AFSCME*

| | controls added | | | |
|---|---|---|---|---|
| | none | economic | +political | +time trends |
| coverage | 0.012 | 0.009 | 0.009 | -0.008 |
| | (0.010) | (0.011) | (0.013) | (0.014) |
| lcoverage | 0.092** | 0.079* | 0.061 | -0.025 |
| | (0.042) | (0.043) | (0.054) | (0.058) |
| freeride | 0.007 | -0.002 | -0.012 | 0.001 |
| | (0.017) | (0.016) | (0.015) | (0.013) |
| udiff | 0.064** | 0.069*** | 0.073*** | 0.054* |
| | (0.025) | (0.024) | (0.027) | (0.029) |
| luwage | 0.029 | 0.023 | 0.021 | 0.025 |
| | (0.020) | (0.018) | (0.019) | (0.019) |
| lnuwage | -0.005 | -0.014 | -0.022 | -0.013 |
| | (0.021) | (0.018) | (0.019) | (0.020) |
| strikehazard | -127.715 | -138.765 | -153.022 | -137.303 |
| | (109.121) | (113.035) | (122.751) | (121.519) |

Robust standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Note: regression sample includes ‘Always RTW’ states,
excludes RTW Switchers.

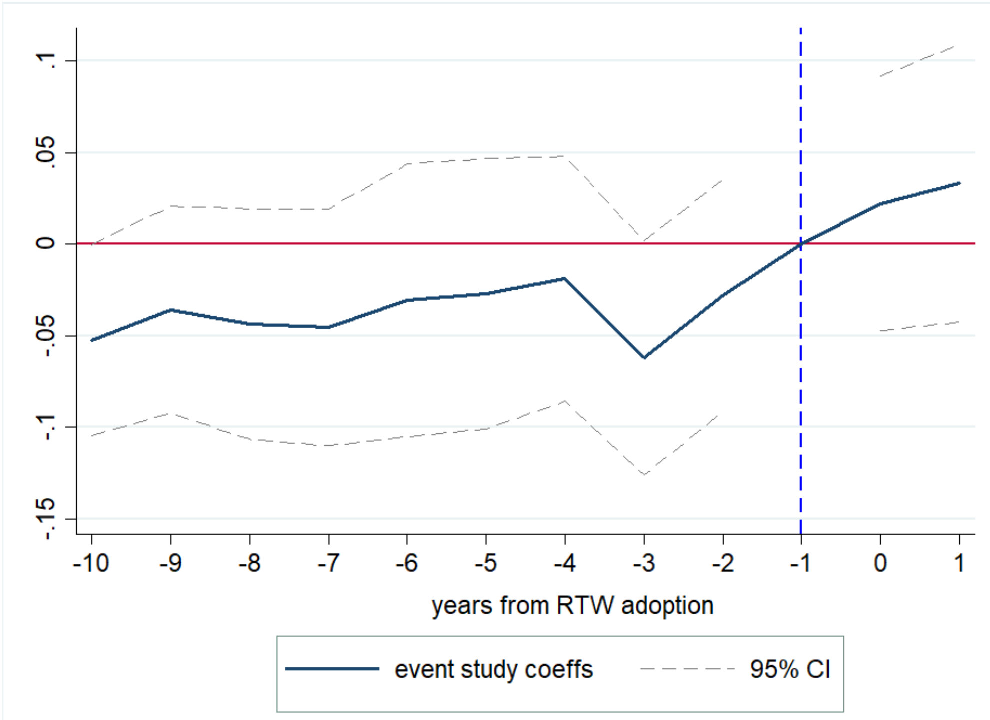



Figure 3.4: Event study of *Janus v. AFSCME*'s effect on public-sector union wage differentials

than 10 percent, and an increase by three out of 100 covered employees, respectively); in the latter case, a RTW effect is almost non-existent; the full-sample results are halfway between these two poles. Part of this contrast may have to do with the shortness of exposure for the states treated in 2018, and part of it may be due to confounding by other state-level policies targeting public sector unions. Time may tell which of these is more important.

Second, union wage differentials increased significantly (both economically and statistically) in the *Janus*-affected states. This result suggests a behavioral response of unions more strongly than does the private sector result: the small and insignificant decline in union coverage rules out compositional change as an explanation for this increase in udiff. Indeed, an increase in udiff driven by heightened union militancy could even explain the absence of an RTW effect on coverage and freeride for the *Janus*-affected states. Figure 3.4 is an event study of the effect of *Janus* on udiff, which illustrates the year-to-year movements of udiff in the *Janus*-affected states relative to the Always-RTW states. An unusual uptick in the event-study coefficient at year $\tau = -1$ is apparent, but this is unsurprising given that the *Janus v. AFSCME* decision was broadly anticipated by observers at least a year in advance.[10] This implies that public sector unions in the original treatment states could, in theory, have reduced the damage of RTW legislation by going on the offensive and winning better contracts. However, they were probably limited in their ability to do so owing to demoralization from previous defeats (such as in Wisconsin 2011) and the resulting restrictions on collective bargaining rights.

## 3.6 Conclusions

Right-to-Work laws are often portrayed as a mortal threat to organized labor in the United States. In academic research as well as in policy debates, it is broadly assumed that RTW laws would harm unions specifically by encouraging free-riding, i.e. by inducing existing union members to succumb to their individualistic instincts. This Free-rider Hypothesis arguably rests on a rather bleak view of unionized workers' collective agency.

This chapter tested the substance of such claims by empirically evaluating the impacts of recently enacted RTW laws on unionization rates, non-membership among covered employees, and other outcomes. Consistent with conventional claims, I find evidence that all else equal, RTW legislation reduces union coverage by more than 10 percent in the private sector; public sector union coverage appears to be similarly affected, but this may be due as much to other policies targeting public sector unions in the newly-RTW states as to the laws themselves. In contrast to the Free-rider Hypothesis, I find non-membership (or free-riding) to increase very little under RTW regimes. This suggests that union attrition under RTW laws may be due more to loss of agency fee revenues affecting unions with already-high shares of nonmembers,

[10] And not just by the American Federation of State, County, and Municipal Employees (the defendant in the case): see, for example, this statement dated May 2017 by the American Federation of Teachers. `https://www.uft.org/news/union-resolutions/resolution-regarding-janus-v-afscme-case`

rather than to loss of existing members. An autopsy of unions that perish under RTW could shed some light on the matter, but this is unfortunately impossible without union-level data.

I also find evidence that RTW laws increased union wage differentials in the private sector by more than two percentage points, and that the 2018 U.S. Supreme Court decision in *Janus v. AFSCME* increased public sector union wage differentials in affected states by more than five percentage points. The *Janus* decision also appears to have had little to no effect on public sector union coverage, at least over a two-year period. These findings suggest that unions respond to RTW laws by bargaining more aggressively to deliver for their constituents, in an attempt to cement their loyalty. In the extreme, this may imply that RTW laws are even beneficial for unionized workers, although life could be made more difficult for union officials compelled to champion workers' interests more zealously. Without observing individual unions making such behavioral adjustments, however, it is not possible to come to definitive conclusions. Future work must investigate whether and how union behavior changes under RTW laws using establishment-level data.

# Bibliography


Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, 105(490):493–505.

Admati, A. R. and Perry, M. (1987). Strategic delay in bargaining. *The Review of Economic Studies*, 54:345.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The economic journal*, 100(401):464–477.

Ashenfelter, O. and Johnson, G. E. (1969). Bargaining theory, trade unions, and industrial strike activity. *The American economic review*, 59(1):35–49.

Atherton, W. N. (1973). *Theory of union bargaining goals*. Princeton University Press.

Bennett, J. T. (1991). Private sector unions: The myth of decline. *Journal of Labor Research*, 12(1):1–12.

Bernstein, I. (2010). *The turbulent years: A history of the American worker, 1933-1941*. Haymarket Books.

Berry, W. D., Ringquist, E. J., Fording, R. C., and Hanson, R. L. (1998). Measuring citizen and government ideology in the american states, 1960-93. *American Journal of Political Science*, pages 327–348.

Blair, D. H. and Crawford, D. L. (1984). Labor union objectives and collective bargaining. *The Quarterly Journal of Economics*, 99(3):547–566.

Blair, P. N. (1974). Union security agreements in public employment. *Cornell L. Rev.*, 60:183.

Blanc, E. (2019). *Red state revolt: The teachers' strike wave and working-class politics*. Verso Books.

Blanchflower, D. G. and Bryson, A. (2004). What effect do unions have on wages now and would freeman and medoff be surprised? *Journal of Labor Research*, 25(3):383–414.

Booth, A. (1984). A public choice model of trade union behaviour and membership. *The Economic Journal*, 94(376):883–898.

Brecher, J. (2014). *Strike!* PM Press.

Brenner, R. (2010). The political economy of the rank-and-file rebellion. In Brenner, A., Brenner, R., and Winslow, C., editors, *Rebel rank and file: labor militancy and revolt from below during the long 1970s*, pages 37–74. Verso.

Brody, D. (1993). *Workers in industrial America: essays on the twentieth century struggle*. Oxford University Press, USA.

Cantor, N. L. (1983). Uses and abuses of the agency shop. *Notre Dame L. Rev.*, 59:61.

Card, D. (2001). The effect of unions on wage inequality in the us labor market. *ILR Review*, 54(2):296–315.

Card, D. and Olson, C. A. (1995). Bargaining power, strike durations, and wage outcomes: An analysis of strikes in the 1880s. *Journal of Labor Economics*, 13(1):32–61.

Carruth, A., Oswald, A. J., and Findlay, L. (1986). A test of a model of trade union behaviour: the coal and steel industries in britain. *Oxford Bulletin of Economics and Statistics*, 48(1):1–18.

Cavallo, E., Galiani, S., Noy, I., and Pantano, J. (2013). Catastrophic natural disasters and economic growth. *Review of Economics and Statistics*, 95(5):1549–1561.

Chaison, G. (2012). *The new collective bargaining*. Springer Science & Business Media.

Cho, I.-K. (1990). Uncertainty and delay in bargaining. *The Review of Economic Studies*, 57(4):575–595.

Clark, D. J. and Riis, C. (1998). Contest success functions: an extension. *Economic Theory*, 11(1):201–204.

Cliff, T. and Gluckstein, D. (1986). *Marxism and trade union struggle: The general strike of 1926*. Bookmarks.

Coase, R. H. (1972). Durability and monopoly. *The Journal of Law and Economics*, 15(1):143–149.

Cogen, C. (1954). The developing law is joining the union required in the taft-hartley union shop? *Labor Law Journal*, 5(10):659.

Craft, J. A., Labovitz, T., and Abboushi, S. (1985). Concession bargaining and unions: Impacts and implications. *Journal of Labor Research*, 6(2):167–180.

Cramton, P. C. (1992). Strategic delay in bargaining with two-sided uncertainty. *The Review of Economic Studies*, 59(1):205–225.

Currie, J. and Ferrie, J. (2000). The law and labor strife in the united states, 1881-1894. *Journal of Economic History*, pages 42–66.

Dark, T. E. (2000). Labor and the democratic party: A report on the 1998 elections. *Journal of Labor Research*, 21(4):627–640.

Darlington, R. (2008). *Radical unionism: the rise and fall of revolutionary syndicalism*. Haymarket books.

Davis, J. C. and Huston, J. H. (1993). Right-to-work laws and free riding. *Economic Inquiry*, 31(1):52–58.

Davis, J. C. and Huston, J. H. (1995). Right-to-work laws and union density: New evidence from micro data. *Journal of Labor Research*, 16(2):223–229.

Davis, M. (1999). *Prisoners of the American dream: Politics and economy in the history of the US working class*. Verso.

Devinatz, V. G. (2015). Right-to-work laws, the southernization of us labor relations and the us trade union movement's decline. *Labor Studies Journal*, 40(4).

DiNardo, J. and Lemieux, T. (1997). Diverging male wage inequality in the united states and canada, 1981–1988: Do institutions explain the difference? *ILR Review*, 50(4):629–651.

Dixon, M. (2007). Limiting labor: Business political mobilization and union setback in the states. *Journal of Policy History*, 19(3):313–344.

Dunlop, J. T. (1944). *Wage determination under trade unions*. New York: Macmillan Co.

Eidlin, B. (2019). Labor unions and movements. In *The Oxford Handbook of Karl Marx*, page 403. Oxford University Press.

Ellwood, D. T. and Fine, G. (1987). The impact of right-to-work laws on union organizing. *Journal of Political Economy*, 95(2):250–273.

Eren, O. and Ozbeklik, S. (2016). What do right-to-work laws do? evidence from a synthetic control method analysis. *Journal of Policy Analysis and Management*, 35(1):173–194.

Farber, H. S. (1984). Right-to-work laws and the extent of unionization. *Journal of labor economics*, 3:319–352.

Farber, H. S. (1986). The analysis of union behavior. *Handbook of labor economics*, 2:1039–1089.

Farber, H. S., Herbst, D., Kuziemko, I., and Naidu, S. (2018). Unions and inequality over the twentieth century: New evidence from survey data. Technical report, National Bureau of Economic Research.

Farhang, S. and Katznelson, I. (2005). The southern imposition: Congress and labor in the new deal and fair deal. *Studies in American Political Development*, 19(1):1.

Feigenbaum, J., Hertel-Fernandez, A., and Williamson, V. (2018). From the bargaining table to the ballot box: Political effects of right to work laws. Technical report, National Bureau of Economic Research.

Fernandez, R. and Glazer, J. (1991). Striking for a bargain between two completely informed agents. *The American Economic Review*, pages 240–252.

Freeman, R. B. and Medoff, J. L. (1984). What do unions do. *Indus. & Lab. Rel. Rev.*, 38:244.

Friedman, G. (2007). *Reigniting the Labor Movement: Restoring means to ends in a democratic Labor Movement*. Routledge.

Friedman, T. J. (2008). Exploiting the north-south differential: Corporate power, southern politics, and the decline of organized labor after world war ii. *The Journal of American History*, 95(2):323–348.

Fudenberg, D. and Tirole, J. (1983). Sequential bargaining with incomplete information. *The Review of Economic Studies*, 50(2):221–247.

Fudenberg, D. and Tirole, J. (1986). A theory of exit in duopoly. *Econometrica: Journal of the Econometric Society*, pages 943–960.

Galiani, S. and Quistorff, B. (2017). The synth_runner package: Utilities to automate synthetic control estimation using synth. *The Stata Journal*, 17(4):834–849.

Garfinkel, M. R. and Skaperdas, S. (2000). Conflict without misperceptions or incomplete information: how the future matters. *Journal of conflict resolution*, 44(6):793–807.

Goldfield, M. (1989). Worker insurgency, radical organization, and new deal labor legislation. *The American Political Science Review*, pages 1257–1282.

Goldin, C. (1994). Labor markets in the twentieth century. *NBER Working Paper*, (h0058).

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Technical report, National Bureau of Economic Research.

Gramm, C. L. (1986). The determinants of strike incidence and severity: A micro-level study. *ILR Review*, 39(3):361–376.

Griffin, L. J., Wallace, M. E., and Rubin, B. A. (1986). Capitalist resistance to the organization of labor before the new deal: Why? how? success? *American Sociological Review*, pages 147–167.

Grossman, J. (1975). The coal strike of 1902-turning point in us policy. *Monthly Labor Review*, 98:21.

Gul, F., Sonnenschein, H., and Wilson, R. (1986). Foundations of dynamic monopoly and the coase conjecture. *Journal of Economic Theory*, 39(1):155–190.

Hart, O. (1989). Bargaining and strikes. *The Quarterly Journal of Economics*, 104(1):25–43.

Hayes, B. (1984). Unions and strikes with asymmetric information. *Journal of Labor Economics*, 2(1):57–83.

Hinich, M. J. (1981). Voting as an act of contribution. *Public choice*, 36(1):135–140.

Hirsch, B. T. and Schumacher, E. J. (2004). Match bias in wage gap estimates due to earnings imputation. *Journal of labor economics*, 22(3):689–722.

Holmes, T. J. (1998). The effect of state policies on the location of manufacturing: Evidence from state borders. *Journal of political Economy*, 106(4):667–705.

Hundley, G. (1988). Who joins unions in the public sector? the effects of individual characteristics and the law. *Journal of Labor Research*, 9(4):301–323.

Hundley, G. (1993). Collective bargaining coverage of union members and nonmembers in the public sector. *Industrial Relations: A Journal of Economy and Society*, 32(1):72–93.

Hunt, J. C. and White, R. A. (1983). The effects of right-to-work legislation on union outcomes: Additional evidence. *Journal of Labor Research*, 4(1):47–63.

Hyman, R. (1975). *Marxism and the sociology of trade unionism*. Pluto Press.

Ichniowski, C. and Zax, J. S. (1991). Right-to-work laws, free riders, and unionization in the local public sector. *Journal of Labor Economics*, 9(3):255–275.

Jacobs, J. B. (2007). *Mobsters, unions, and feds: The mafia and the American labor movement*. NYU Press.

Jorgenson, D. W., Ho, M. S., and Stiroh, K. J. (2008). A retrospective look at the us productivity growth resurgence. *Journal of Economic perspectives*, 22(1):3–24.

Kalenkoski, C. M. and Lacombe, D. J. (2008). Effects of minimum wages on youth employment: the importance of accounting for spatial correlation. *Journal of Labor Research*, 29(4):303–317.

Kelly, J. (2012). *Rethinking industrial relations: Mobilisation, collectivism and long waves*. Routledge.

Kennan, J. (1986). The economics of strikes. *Handbook of labor economics*, 2:1091–1137.

Kennan, J. and Wilson, R. (1989). Strategic bargaining models and interpretation of strike data. *Journal of Applied Econometrics*, 4(S1):S87–S130.

Klare, K. E. (1977). Judicial deradicalization of the wagner act and the origins of modern legal consciousness, 1937-1941. *Minn. L. Rev.*, 62:265.

Klarner, C. (2018). State legislative election returns, 1967-2016. *Harvard Dataverse*, 3.

Lafer, G. (2013). *The legislative attack on American wages and labor standards, 2011-2012*. Economic Policy Institute.

Lapmann, T. (2015). Surprising results from indianan's right-to-work law. *The Buckeye Institute for Public Policy Solutions*.

Lewis, H. G. (1963). *Unionism and relative wages in the United States: an empirical inquiry.* Chicago: University of Chicago Press.

Lichtenstein, N. (2013). *State of the Union: A Century of American Labor-Revised and Expanded Edition*. Princeton University Press.

Lumsden, K. and Petersen, C. (1975). The effect of right-to-work laws on unionization in the united states. *Journal of Political Economy*, 83(6):1237–1248.

Lynd, A. and Lynd, R. S. (2014). *Rank and file: Personal histories by working-class organizers*. Princeton University Press.

Makridis, C. A. (2019). Do right-to-work laws work? evidence on individuals' well-being and economic sentiment. *The Journal of Law and Economics*, 62(4):713–745.

Manzo, F. and Bruno, R. (2017). The impact of "right-to-work" laws on labor market outcomes in three midwest states: Evidence from indiana, michigan, and wisconsin (2010–2016). *Project for Middle Class Renewal*.

Margolis, H. (1984). *Selfishness, altruism, and rationality*. University of Chicago Press.

Martin, D. L. (1980). *An ownership theory of the trade union*. University of California Press.

McBride, M. and Skaperdas, S. (2007). Explaining conflict in low-income countries: incomplete contracting in the shadow of the future. In Gradstein, M. and Konrad, K. A., editors, *Institutions and norms in economic development*, pages 141–161. MIT Press.

McCammon, H. J. (1990). Legal limits on labor militancy: Us labor law and the right to strike since the new deal. *Social Problems*, 37(2):206–229.

McCartin, J. A. (1997). *Labor's Great War: The Struggle for Industrial Democracy and the Origins of Modern American Labor Relations, 1912-1921*. UNC Press Books.

McDonald, I. M. and Solow, R. M. (1981). Wage bargaining and employment. *The American Economic Review*, 71(5):896–908.

McNally, D. (2011). *Global slump: The economics and politics of crisis and resistance*. pm Press.

Michels, R. (1915). *Political parties: A sociological study of the oligarchical tendencies of modern democracy*. Hearst's International Library Company.

Mills, C. W. (1948). *The new men of power: America's labor leaders*. Harcourt, Brace and Company.

Mitchell, M. S. (1978). Public sector union security: The impact of" abood". *Labor Law Journal*, 29(11):697.

Montgomery, D. (1980). *Workers' control in America: Studies in the history of work, technology, and labor struggles*. Cambridge University Press.

Moody, K. (1988). *An injury to all: The decline of American unionism*. Verso.

Moody, K. (2010). Understanding the rank-and-file rebellion in the long 1970s. In Brenner, A., Brenner, R., and Winslow, C., editors, *Rebel rank and file: labor militancy and revolt from below during the long 1970s*, pages 105–146. Verso.

Moody, K. (2014). *In Solidarity: Essays on Working-Class Organization and Strategy in the United States*. Haymarket Books.

Moore, W. J. (1980). Membership and wage impact of right-to-work laws. *Journal of Labor Research*, 1(2):349–368.

Moore, W. J. (1998). The determinants and effects of right-to-work laws: A review of the recent literature. *Journal of Labor Research*, 19(3):445–469.

Moore, W. J., Dunlevy, J. A., and Newman, R. J. (1986). Do right to work laws matter? comment. *Southern Economic Journal*, 53(2):515–524.

Moore, W. J. and Newman, R. J. (1985). The effects of right-to-work laws: A review of the literature. *ILR Review*, 38(4):571–585.

Murphy, K. J. (1992). Determinants of contract duration in collective bargaining agreements. *ILR Review*, 45(2):352–365.

Nack, D., Childers, M., Kulwiec, A., and Ibarra, A. (2020). The recent evolution of wisconsin public worker unionism since act 10. *Labor Studies Journal*, 45(2):147–165.

Naidu, S. and Yuchtman, N. (2016). Labor market institutions in the gilded age of american economic history. Technical report, National Bureau of Economic Research.

Nalebuff, B. and Riley, J. (1985). Asymmetric equilibria in the war of attrition. *Journal of Theoretical Biology*, 113(3):517–527.

Oswald, A. J. (1992). The microeconomic theory of the trade union. In Garonna, P., Mori, P., and Tadeschi, P., editors, *Economic Models of Trade Unions*, pages 3–24. Springer.

Perry, L. J. and Wilson, P. J. (2007). Convergence of work stoppages: A global perspective. *Journal of World-Systems Research*, pages 198–219.

Reed, W. R. (2003). How right-to-work laws affect wages. *Journal of Labor Research*, 24(4):713–730.

Reich, R. B. (2016). *Saving capitalism: For the many, not the few*. Vintage.

Romer, C. D. (1989). The prewar business cycle reconsidered: New estimates of gross national product, 1869-1908. *Journal of Political Economy*, 97(1):1–37.

Ross, A. M. (1956). *Trade union wage policy*. Univ of California Press.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica: Journal of the Econometric Society*, pages 97–109.

Sanes, M. and Schmitt, J. (2014). *Regulation of public sector collective bargaining in the states*. Center for Economic and Policy Research Washington, DC.

Schaller, Z. (2019). The decline of u.s. labor unions: the role of trade and sectoral shifts. *Job market paper*.

Schaller, Z. and Skaperdas, S. (2020). Bargaining and conflict with up-front investments: How power asymmetries matter. *Journal of Economic Behavior & Organization*, 176:212–225.

Schumacher, E. J. (1999). What explains wage differences between union members and covered nonmembers? *Southern Economic Journal*, pages 493–512.

Selfa, L. (2012). *The Democrats: A Critical History (Updated edition)*. Haymarket books.

Silver, B. J. (2003). *Forces of labor: workers' movements and globalization since 1870*. Cambridge University Press.

Skaperdas, S. (1996). Contest success functions. *Economic theory*, 7(2):283–290.

Skaperdas, S. and Vaidya, S. (2012). Persuasion as a contest. *Economic theory*, 51(2):465–486.

Slaughter, J. (1983). Concessions bargaining in auto. *Labor Research Review*, 1:70–72.

Smith, J. M. (1974). The theory of games and the evolution of animal conflicts. *Journal of theoretical biology*, 47(1):209–221.

Sobel, J. and Takahashi, I. (1983). A multistage model of bargaining. *The Review of Economic Studies*, 50(3):411–426.

Sobel, R. S. (1995). Empirical evidence on the union free-rider problem: Do right-to-work laws matter? *Journal of Labor Research*, 16(3):347–365.

Stevans, L. K. (2009). The effect of endogenous right-to-work laws on business and economic conditions in the united states: A multivariate approach. *Review of Law & Economics*, 5(1):595–614.

Stokey, N. L. (1981). Rational expectations and durable goods pricing. *The Bell Journal of Economics*, pages 112–128.

Stricker, F. (1983). Affluence for whom?—another look at prosperity and the working classes in the 1920s. *Labor History*, 24(1):5–33.

Sugden, R. (1984). Reciprocity: the supply of public goods through voluntary contributions. *The Economic Journal*, 94(376):772–787.

Swindle, S. R. (1984). Drawing lines and defining remedies: The impact of ellis v. brotherhood of railway, airline and steamship clerks on the first amendment rights of dissident employees. *BYU L. Rev.*, page 647.

Tracy, J. S. (1987). An empirical test of an asymmetric information model of strikes. *Journal of Labor Economics*, 5(2):149–173.

Tullock, G. (1980). Efficient rent-seeking. In Buchanan, J., Tollinson, R., and Tullock, G., editors, *Toward a Theory of the Rent-Seeking Society*, pages 97–112. College Station: Texas AM Press.

Voss, K. (1993). *The making of American exceptionalism: The knights of labor and class formation in the nineteenth century*. Ithaca, NY: Cornell University Press.

Voss, K. (2010). Democratic dilemmas: union democracy and union renewal. *Transfer: European Review of Labour and Research*, 16(3):369–382.

Vroman, S. B. (1989). A longitudinal analysis of strike activity in us manufacturing: 1957-1984. *The American Economic Review*, 79(4):816–826.

Wallace, M., Rubin, B. A., and Smith, B. T. (1988). American labor law: Its impact on working-class militancy, 1901-1980. *Social Science History*, 12(1):1–29.

Webb, S. and Webb, B. (1920). *The history of trade unionism*. Longmans, Green.

Wessels, W. J. (1981). Economic effects of right to work laws. *Journal of Labor Research*, 2(1):55–75.

Western, B. and Rosenfeld, J. (2011). Unions, norms, and the rise in us wage inequality. *American Sociological Review*, 76(4):513–537.

Winslow, C. (2010a). *Labor's Civil War in California: The NUHW Healthcare Workers' Rebellion*. PM Press.

Winslow, C. (2010b). Overview: The rebellion from below 1965–81. In Brenner, A., Brenner, R., and Winslow, C., editors, *Rebel rank and file: labor militancy and revolt from below during the long 1970s*, pages 1–35. Verso.

Zinn, H. (2015). *A people's history of the United States: 1492-present*. Routledge.

# Appendix A

# Chapter One Appendix

## A.1 An illustrative model of collective action success/failure

Suppose there is an economic surplus of size 1 that workers can acquire by winning a strike. Individual workers are a point mass along a unit interval, and are indexed by $i \in [0, 1]$. The probability of winning a strike depends linearly on the share of workers participating, or $p \in [0, 1]$. Workers' expected wage gain conditional on striking is therefore equal to the expected participation rate $E(p)$. Worker utility assumes the following form:

$$EU_i = \begin{cases} i \cdot E(p) - \frac{c}{E(p)} + E(p) & \text{if participate} \\ E(p) & \text{if stay out} \end{cases}$$

In other words, worker $i$ participates if $i \cdot E(p) - \frac{c}{E(p)} \geq 0$. The $i \cdot E(p)$ term represents the level of gratification worker $i$ derives from contributing to an expected gain in her colleagues' well-being: it captures a pro-social (other-regarding) component of worker utility that is assumed to vary from person to person (e.g. the worker indexed by $i = 0$ is a purely self-interested free-rider). $c$ captures the costs to participating in a strike, such as the risk of getting fired or

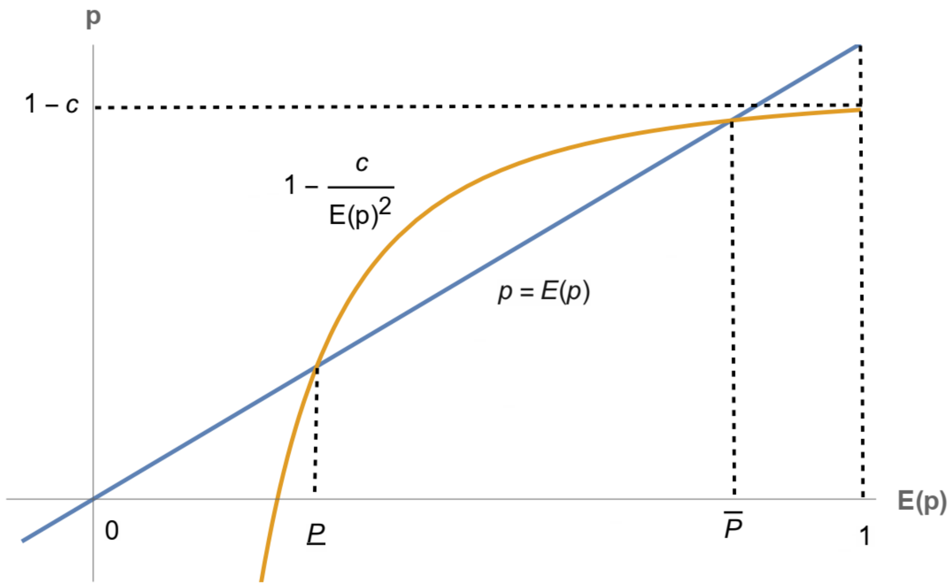



Figure A.1: Two possible equilibria in strike participation level

permanently replaced. The cost to individual strike participants is assumed to be declining in the rate of participation, because there is in general a limit to the number of strikers an employer can replace or victimize without causing serious disruption to business, and hence a lower chance of victimization when strikers have numbers on their side.

It follows that all workers with $i$ such that $i \cdot E(p) \geq \frac{c}{E(p)}$ will participate. The pivotal participant is indexed by $i^* = \frac{c}{E(p)^2}$ (which is less than 1 if $E(p) \geq \sqrt{c}$), so that the actual participation rate will be:

$$p = 1 - i^* = 1 - \frac{c}{E(p)^2}$$

In equilibrium, expectations will have to converge to actual participation, so that $p = E(p)$. Plotting these two equations together in Figure A.1, we see that two equilibrium participation rates are possible, one low ($\underline{p}$) and the other high ($\overline{p}$).

Only $\overline{p}$ is stable. This is the point that $p$ will converge to if initial expectation $E_0(p)$ is even slightly higher than $p$; for any lower initial expectation, the putative strike quickly peters out.

This simple model suggests a number of ways in which the outcome of a strike in one period can alter the union's winning probability $\lambda$ in future periods.

For example, suppose initial expectation $E_0(p)$ is a random variable with some distribution $G(.)$ that is common to all workers and is realized at the beginning of an already-launched strike. Then before the strike is launched, the probability of winning is $\lambda(G) = (1 - G(\underline{p}))\overline{p}$. It is not too hard to imagine that a strike victory may shift $G(.)$ to the right (and $\lambda(G)$ up): in all likelihood a strike won because a high share ($\overline{p}$) of workers turned out, and based on this experience workers adjust upward their expectations on future turnout. Conversely, a weak turnout and the resulting defeat could sow pessimism about future levels of participation, shifting $G(.)$ to the left (and driving $\lambda(G)$ down).

## A.2 Formal derivation of the comparative statics results in proposition 1

Inequality 1.16 defined the condition for the existence of a fighting equilibrium when $\lambda = \underline{\lambda}$ as follows:

$$\varphi\delta\underline{P}(\overline{\lambda} - \underline{\lambda})(\overline{\lambda} + \underline{\lambda} - 1) > \frac{1 - \varphi}{2},$$

$$\text{where} \quad \underline{P} = P(\underline{\lambda}, e^F(\underline{\lambda})) = \frac{\underline{\lambda}(1 + \delta(\overline{\lambda} + \underline{\lambda})(\overline{\lambda} - \underline{\lambda}))}{1 + \delta(\overline{\lambda} - \underline{\lambda})(1 + (\overline{\lambda} + \underline{\lambda} - 1)(2\underline{\lambda} - 1))}.$$

Reparameterizing with bias $\equiv B = \overline{\lambda} + \underline{\lambda} - 1$ and stakes $\equiv D = \overline{\lambda} - \underline{\lambda}$, the fighting equilibrium condition becomes:

$$\frac{\delta BD(B - D + 1)(\delta D(B + 1) + 1)}{\delta D(B(B - D) + 1) + 1} > \frac{1 - \varphi}{\varphi}, \text{s.t.} \quad B - D > -1 \text{ and } B + D < 1.$$

The RHS of the above condition is decreasing in $\varphi$. The LHS is a function of $\delta$, $B$, and $D$, and we check if the partial derivatives are positive over most of the relevant parameter space.

It is not too hard to show that $\frac{\partial LHS}{\partial \delta}$ is always positive as long as $B > 0$:

$$\frac{\partial LHS}{\partial \delta} = \frac{BD(B-D+1)\left[1+2\delta D(1+B)+\delta^2 D^2(1+B(1-D)+B^2(1-D)+B^3))\right]}{\left[1+\delta D(1+B(B-D))\right]^2} > 0$$

It can be checked that every term in both the numerator and denominator of the above expression is positive.

The same is true for $\frac{\partial LHS}{\partial B}$:

$$\frac{\partial LHS}{\partial B} = \delta D\left[1-D+2B+2\delta(D(1-D)+BD(3-D+B))+\delta^2 D^2((1-D)(4B+1)\right.$$
$$\left.+2BD(1-B^2)+B^2(2-D+D^2+B^2))\right] / \left[1+\delta D(1+B(B-D))\right]^2 > 0.$$

$\frac{\partial LHS}{\partial D}$ is more complicated:

$$\frac{\partial LHS}{\partial D} = \delta B\left[\delta^2 B^4 D^2+2\delta^2 B^3 D^2(1-D)+(2D-1)(1+\delta D)^2+\delta B^2 D(2+2\delta D-2\delta D^2+\delta D^3)\right.$$
$$\left.+B(1+4\delta D+2\delta(\delta-1)D^2=2\delta^2 D^3+\delta^2 D^4)\right] / \left[1+\delta D(1+B(B-D))\right]^2.$$

Graphically plotting the above derivative with $B$ and $D$ on the horizontal axes and setting arbitrary values for $\delta$ reveals that $\frac{\partial LHS}{\partial D}$ is positive over most of the $(B, D)$ space, and that it is only slightly negative (close to zero, represented by the blue plane in Figure A.2) in a small subset of the $(B, D)$ space. This result is robust to varying $\delta$ values. All this implies that the comparative static results in proposition 1 are broadly correct: the likelihood of a fighting equilibrium is increasing in $\delta$, bias, and stakes, conditional on bias being positive for the disadvantaged player.

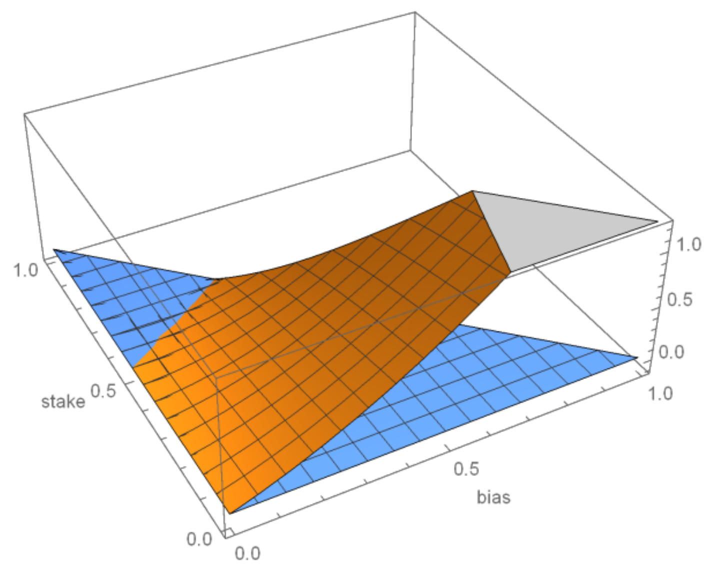



Figure A.2: Plot of $\frac{\partial LHS}{\partial D}$ in $(B, D)$ space, with $\delta = 1.2$

# Appendix B

# Chapter Three Appendix

## B.1 Data Appendix

This section describes in detail how I constructed each of the variables used in Chapter 2. Note that some of the variables described here do not appear in the main text, as they are only used as inputs in the construction of our variables.

### B.1.1 CPS variables

All CPS-based variables except 'popul', 'emp', 'unemp', and 'manuf' have been computed separately for the private and public sector. I have excluded federal government employees in computing the public sector variables. All variables are computed from the CPS Merged Outgoing Rotation Groups (MORG) files released by the National Bureau of Economic Research (NBER), available for download at `https://www.nber.org/research/data/current-population-survey-cps-merged-outgoing-rotation-group-earnings-data`; the Earner Study weights ('earnwt') are used for all aggregation and estimation purposes.

**'popul'**: this is the total population of a state in a given year, used as denominator for per-capita variables. This is obtained by summing the individual weights in each state-year cell and dividing by 12.

**'emp'**: share of persons aged between 16 and 64 who are currently employed (including self-employed) and at work.

**'unemp'**: share of those looking for work or on layoff (unemployed) among a workforce defined as all wage and salary workers plus the unemployed.

**'manuf'**: share of wage and salary workers whose industry code 'dind' (an NBER-created two-digit industry classification code) is between 5 and 28 in years prior to 2000 and 'dind02' is between 5 and 20 for 2000 onward.

**'coverage'**: share of employed (private- or public-sector) wage and salary workers who report being either a union member or covered by a union contract.

**'freeride'**: share of workers covered by union contracts who report not being a union member.

**'uwage'**: weighted average hourly earnings (in nominal dollars) of wage and salary workers covered by union contracts. For workers not paid on an hourly basis, this is usual weekly earning at the main job divided by usual hours worked per week at the main job.

**'nuwage'**: the non-union counterpart to 'uwage'.

**'udiff'**: union wage differential estimated for each state×year cell, or $\hat{\beta}$ in the wage regression:

$$lwage_i = \alpha + \beta union_i + X_i \Gamma + \epsilon_i,$$

where $i$ indexes workers, *lwage* is logged hourly wage, *union* is the union coverage indicator, $X$ is a vector of worker and job attributes (age, age squared, educational attainment, sex, marital status, black, part-time status, whether paid by the hour, dummies for two-digit industry and occupation codes), and $\epsilon$ is an error term. A well-known source of bias in estimating the union wage differential using CPS data is that earnings for individuals who do not report them are imputed based on the earnings of comparable workers without accounting for their unionization status: this results in attenuation bias for estimating the union wage differential (Hirsch and Schumacher, 2004). To address this, I remove all observations with allocated (imputed) earnings for estimating the wage equation.

### B.1.2 strike activity measures

These variables are also disaggregated by sector (private and public). They are based on the US Bureau of Labor Statistics (BLS)'s Work Stoppages tables (Detailed Monthly Listing, 1993-Present) available at `https://www.bls.gov/wsp/` and the US Federal Mediation and Conciliation Service (FMCS)'s work stoppages tables (which had been previously available in Excel format at `https://www.fmcs.gov/resources/documents-and-data/`). The BLS data records all strikes involving more than 1,000 workers; the FMCS data records all smaller strikes. Unlike the BLS tables, the FMCS tables do not classify struck establishments by ownership status (government or private industry). To identify public sector strikes in FMCS records, I looked for employer names containing the strings "school", "unified school district", "USD", "U.S.D.", "university of", "government", "authority", "of education", "state of", "county", and "municipal".

**'counts'**: number of work stoppages in each state that began in a calendar year and lasted no longer than half a year. Strikes that are missing end dates, and all ongoing strikes as of the end of 2019 are excluded, as are strikes classified as "interstate".

**'idled'**: annual number of workers involved in all strikes (that are not excluded in 'counts') in a given state.

**'strikehazard'**: 'idled' times 1,000 divided by the number of unionized workers estimated from CPS data.

**'duration'**: $\frac{1}{N}\Sigma_i n_i \cdot d_i$, where $n_i$ =number of workers idled in strike $i$, $N = \Sigma_i n_i$, and $d_i$=duration of strike $i$ measured in business days. The FMCS did not record $d_i$, so I calculated the total number of days (including weekends) a strike lasted based on its beginning and end dates, and multiplied by $\frac{5}{7}$. Some state-year cells are missing duration values because no strike was observed in those cells: it is difficult to say how long a strike would have lasted had there been one.

### B.1.3 NLRB elections-related variables

These variables are constructed using a dataset on NLRB elections compiled by Zachary Schaller in Schaller (2019), who generously shared it with me for use in this chapter.

**'cert'**: number of union certification elections (identified by a dummy variable for RC petitions) where the percentage of votes in favor ('pct_for') is greater than 0.5.

**'decert'**: number of de-certification elections (RD petitions filed by employees and RM petitions filed by employers) where 'pct_for' is below 0.5.

**'inflow'**: total number of eligible employees in certification elections won by unions.

**'outflow'**: total number of eligible employees in de-certification elections lost by unions.

**'inflow_perK'**: 'inflow' times 1,000 divided by the number of non-union workers.

**'outflow_perK'**: 'outflow' times 1,000 divided by the number of unionized workers.

### B.1.4 Other variables

**'lgdp'**: log of annual state per-capita GDP. Gross nominal state GDP is obtained from the U.S. Bureau of Economic Analysis' Series SAGDP1; this is then divided by 'popul' to generate per-capita GDP.

**'Dgdp'**: first-differenced 'lgdp'.

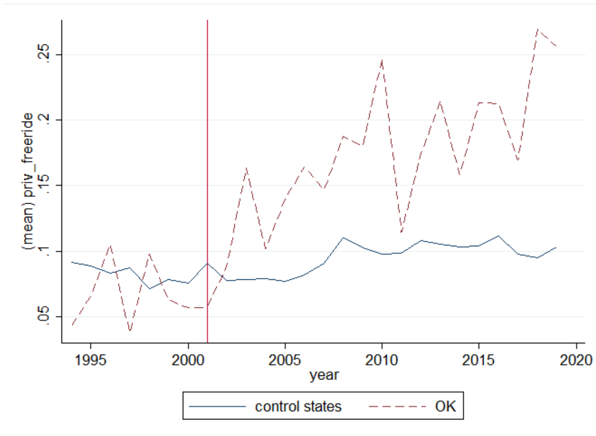
**'house_dem'**: I use the State Legislative Elections, 1971-2016 dataset maintained by the Princeton Gerrymandering Project (available at https://github.com/PrincetonUniversity/historic_state_legislative_election_results) as the primary source for this variable. The number of Democratic candidates who won in their district were counted and divided by the total number of seats. For state-year cells missing this information, I inputted numbers from Wikipedia searches of state-level general election results. This variable is assumed to be unchanging during off-election years.

**'dgov', 'R(D)_legctrl', and 'R(D)_trifecta'**: constructed from state-government partisan composition tables available on the National Conference of State Legislatures website (https://www.ncsl.org/research/about-state-legislatures/partisan-composition.aspx#Timelines).

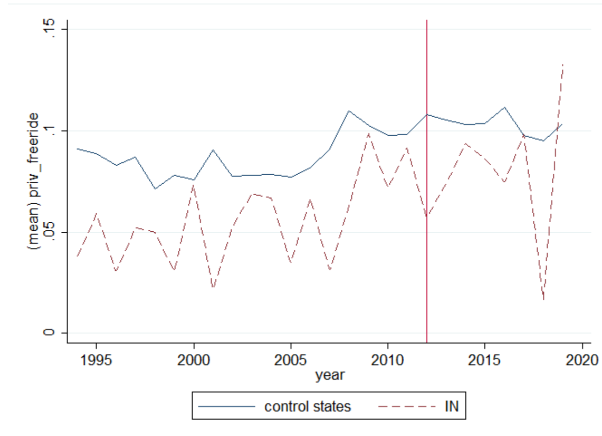
**'ctz_ideo' and 'gtv_ideo'**: for years up to 2016 (for ctz_ideo) and 2017 (for gvt_ideo), these measures are taken directly from an updated version of Berry et al. (1998)'s series available at https://rcfording.com/state-ideology-data/. The measures are on a scale of zero to 100, zero being the most conservative and 100 the most liberal. Berry et al. (1998) constructed them using data on election returns and subjective ratings of the ideologies of individual members of U.S. Congress representing different states. These ratings are furnished by two interest groups, the AFL-CIO Committee on Political Education (COPE) and Americans for Democratic Action (ADA). I impute ctz_ideo (_ideo) values for 2017

(2018) and onward using predicted values generated from a regression of observed values on their third (second) lags, state fixed-effects, emp, lgdp, manuf, and the full set of political variables that I have for all years.

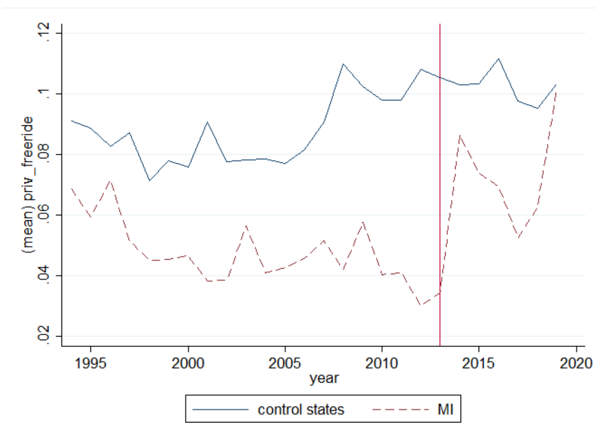
## B.2 Figures and Tables



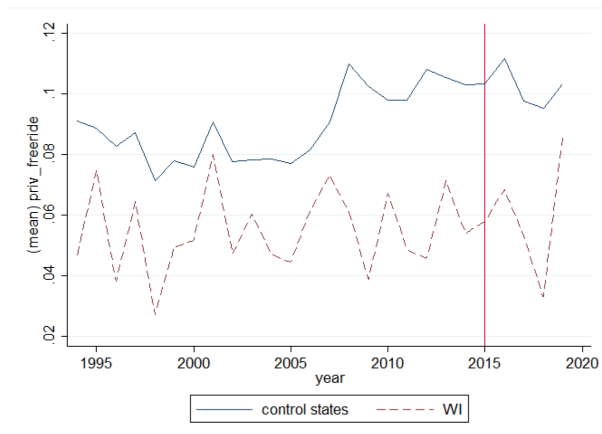

(a) Oklahoma



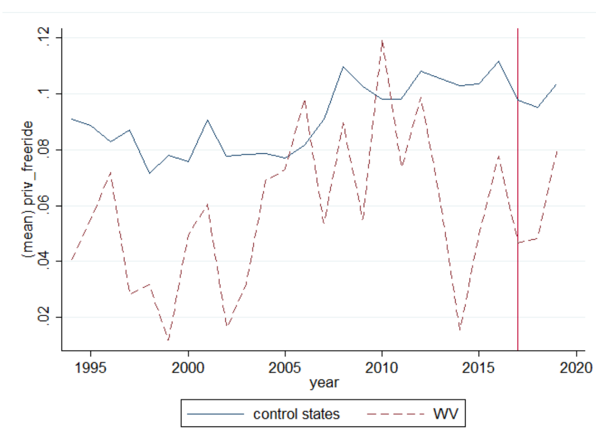

(b) Indiana



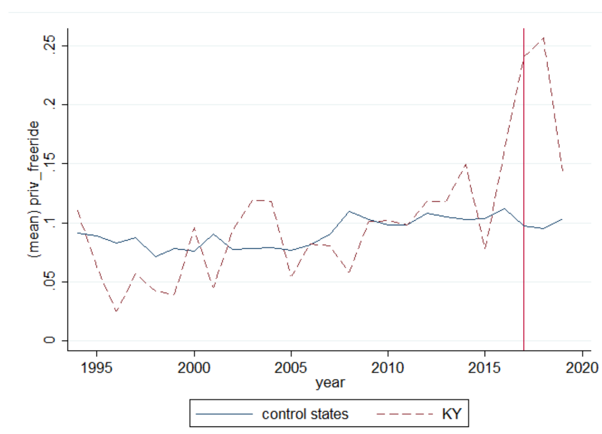

(c) Michigan




(d) Wisconsin

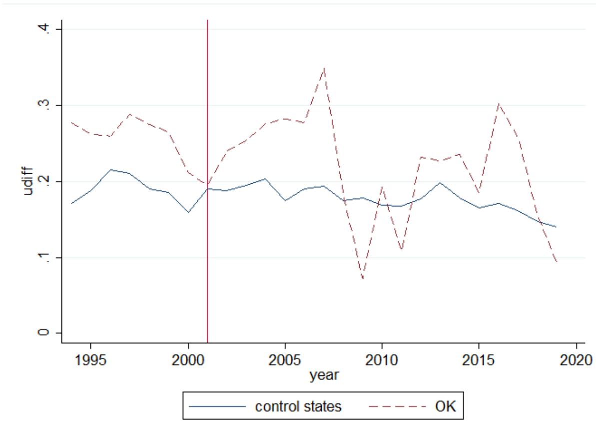



(e) West Virginia

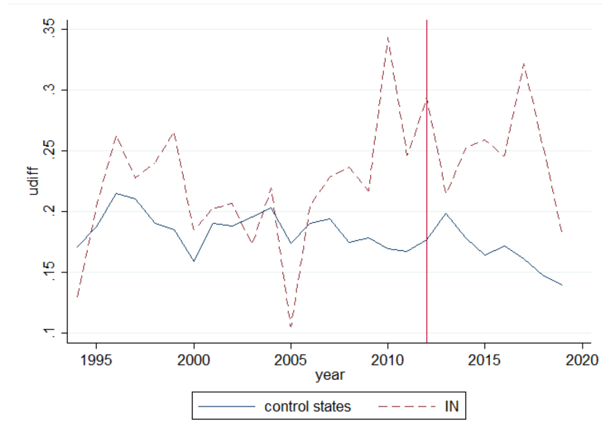



(f) Kentucky

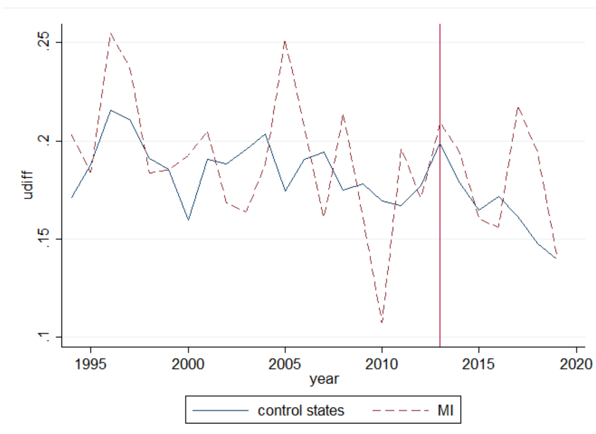
Figure B.1: Raw trends in private-sector free-rider share



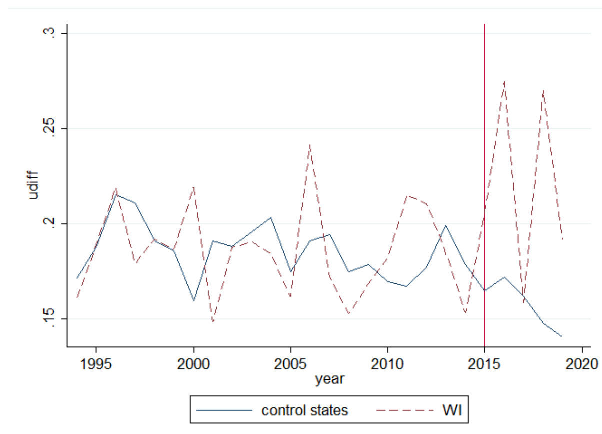

(a) Oklahoma



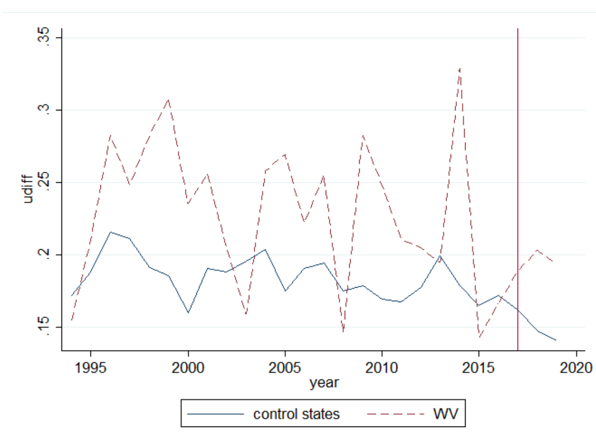

(b) Indiana



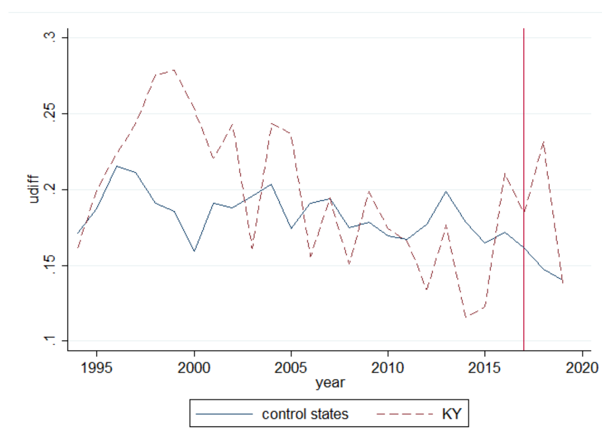

(c) Michigan




(d) Wisconsin

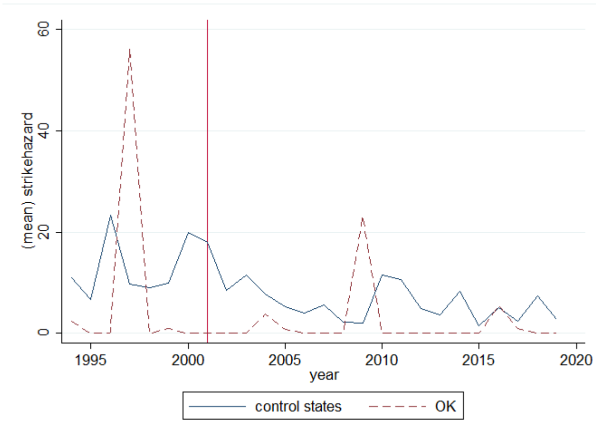



(e) West Virginia

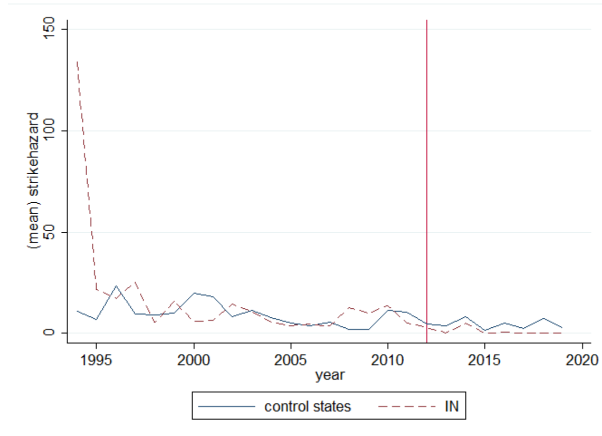



(f) Kentucky

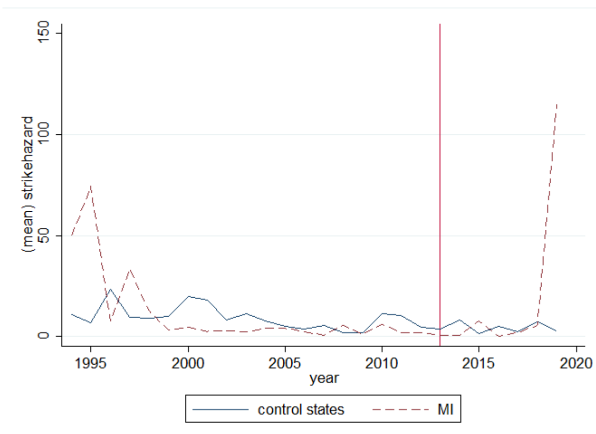
Figure B.2: Raw trends in private-sector union wage differentials



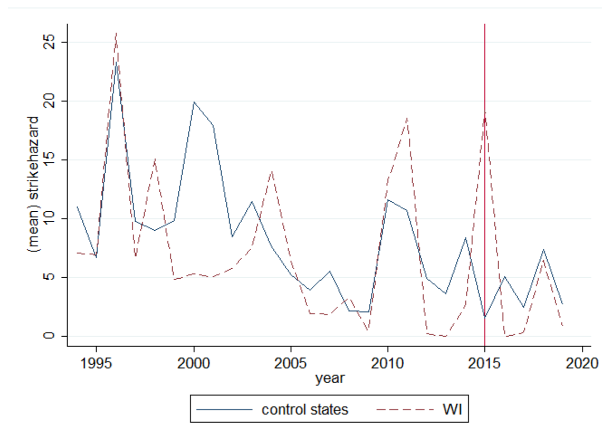

(a) Oklahoma



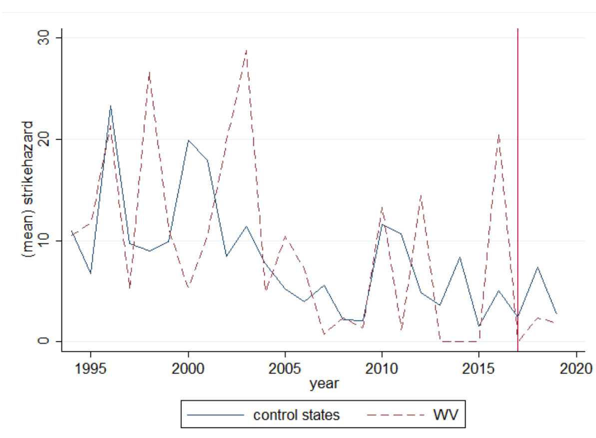

(b) Indiana



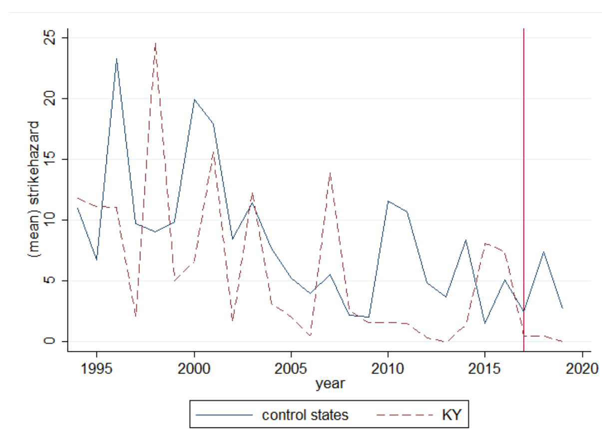

(c) Michigan




(d) Wisconsin

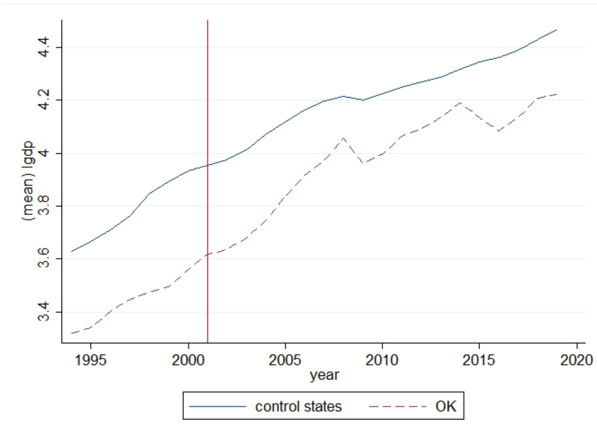



(e) West Virginia

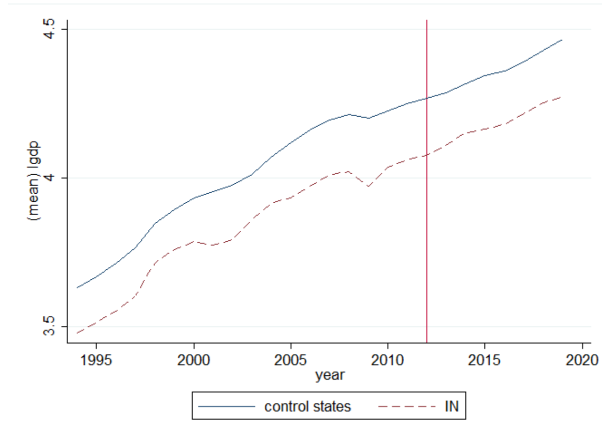



(f) Kentucky

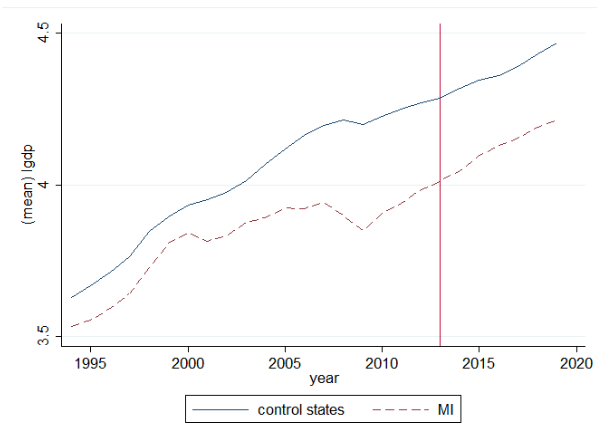
Figure B.3: Raw trends in private-sector strike participants per 1,000



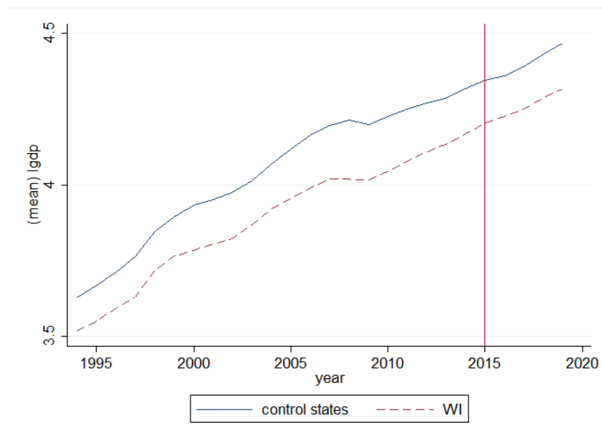

(a) Oklahoma



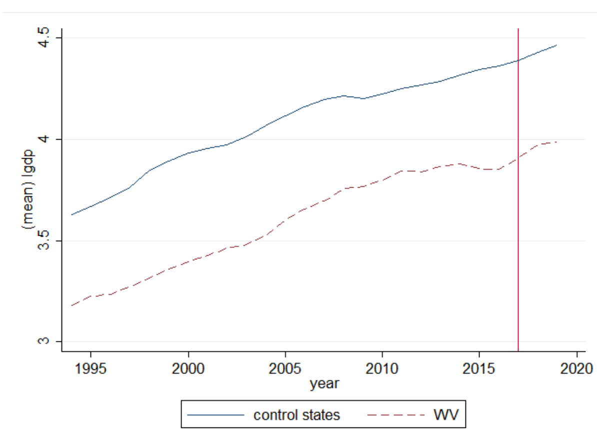

(b) Indiana



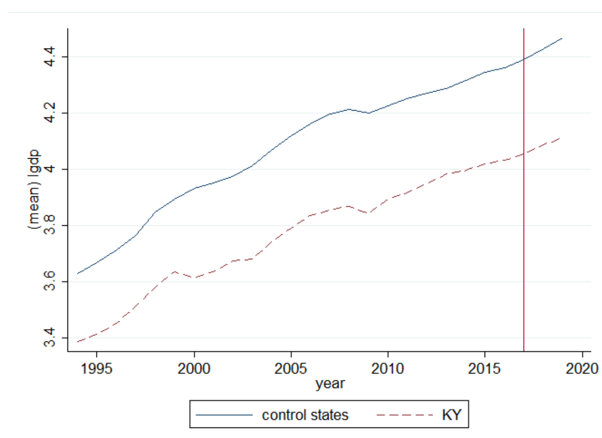

(c) Michigan




(d) Wisconsin

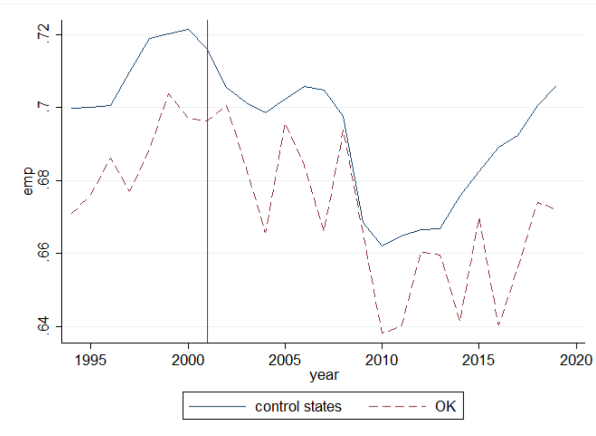



(e) West Virginia



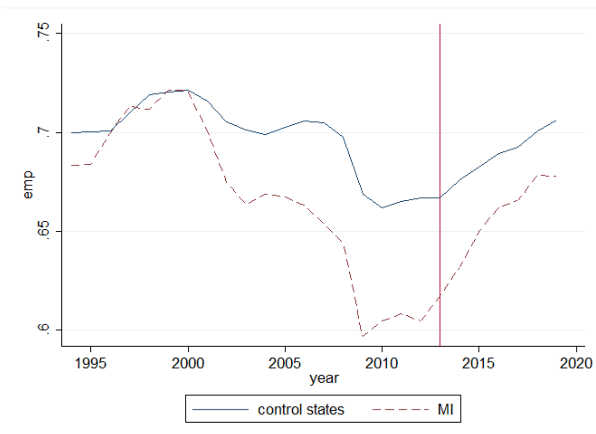

(f) Kentucky

Figure B.4: Raw trends in log of per-capita GDP

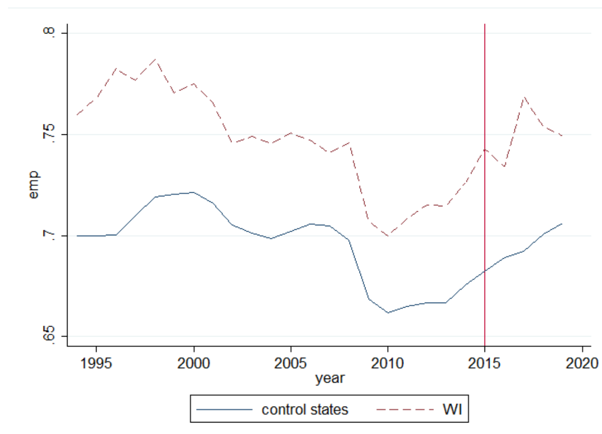



(a) Oklahoma


(b) Indiana

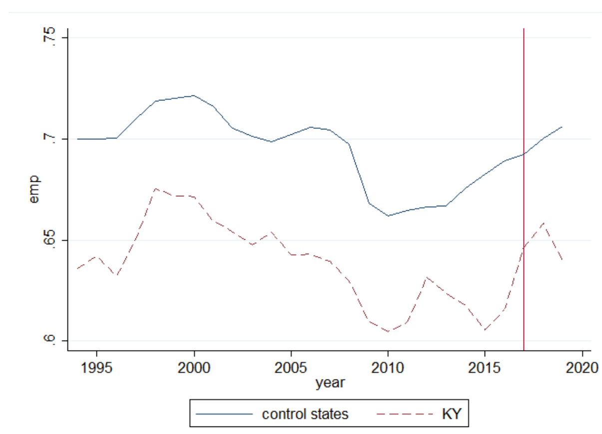



(c) Michigan



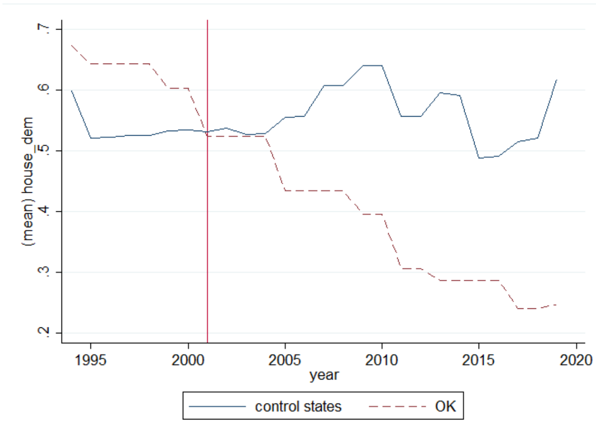

(d) Wisconsin


(e) West Virginia

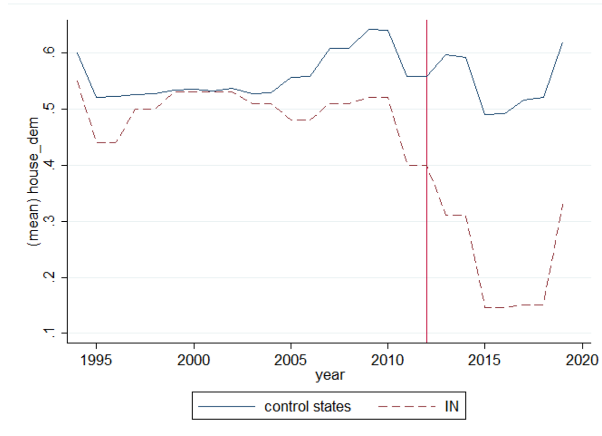



(f) Kentucky

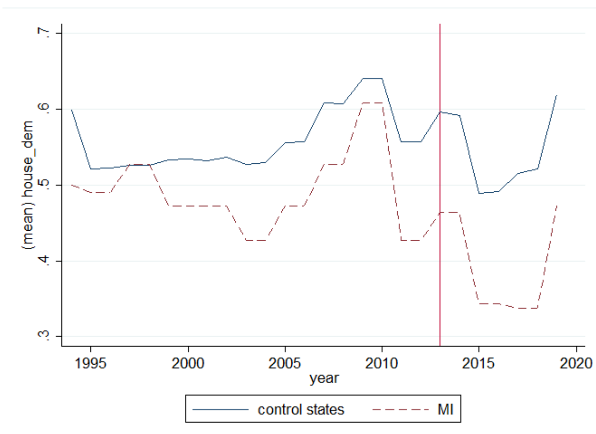
Figure B.5: Raw trends in employment rate



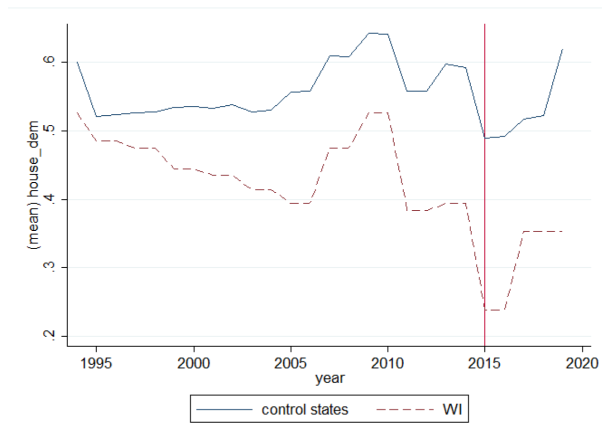

(a) Oklahoma



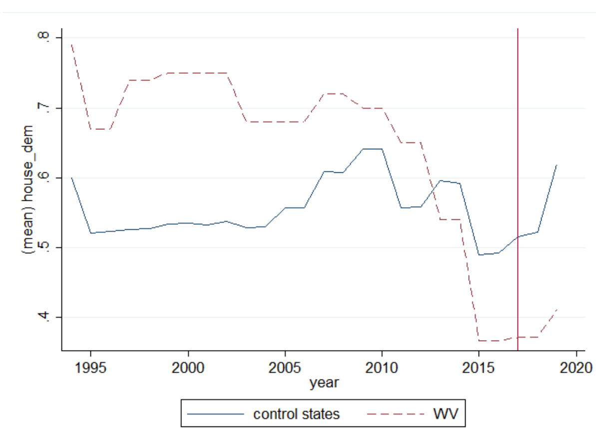

(b) Indiana



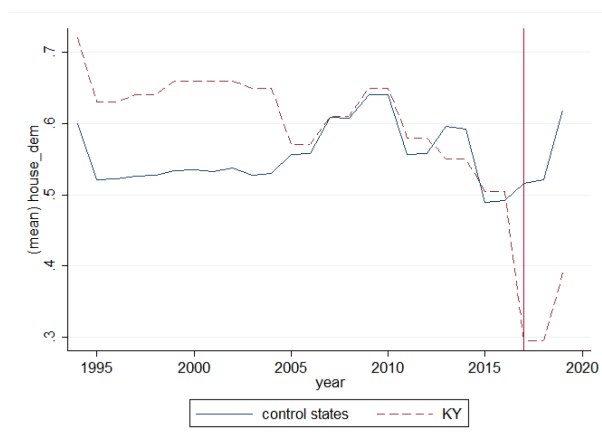

(c) Michigan




(d) Wisconsin




(e) West Virginia




(f) Kentucky

Figure B.6: Raw trends in Democratic share of seats in state lower house

Table B.1: Estimates of effect on union coverage

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | -0.014** | -0.003 | -0.016*** | -0.014** | -0.015*** | -0.009** |
| | (0.007) | (0.005) | (0.006) | (0.006) | (0.005) | (0.004) |
| lgdp | | | 0.033* | | 0.033* | 0.027* |
| | | | (0.017) | | (0.017) | (0.013) |
| Dgdp | | | -0.003 | | -0.015 | 0.003 |
| | | | (0.024) | | (0.028) | (0.021) |
| emp | | | 0.090* | | 0.139*** | 0.071* |
| | | | (0.046) | | (0.040) | (0.039) |
| unemp | | | -0.143** | | -0.068 | -0.040 |
| | | | (0.069) | | (0.062) | (0.054) |
| manuf | | | -0.073 | | -0.003 | 0.011 |
| | | | (0.069) | | (0.061) | (0.055) |
| college | | | -0.096* | | 0.008 | -0.012 |
| | | | (0.052) | | (0.047) | (0.025) |
| house_dem | | | | -0.001 | 0.003 | -0.029*** |
| | | | | (0.013) | (0.012) | (0.008) |
| dgov | | | | -0.007 | -0.005 | -0.004 |
| | | | | (0.006) | (0.004) | (0.003) |
| R_legctrl | | | | -0.001 | -0.002 | 0.001 |
| | | | | (0.004) | (0.003) | (0.002) |
| D_legctrl | | | | -0.007* | -0.006** | 0.000 |
| | | | | (0.004) | (0.003) | (0.002) |
| R_trifecta | | | | 0.001 | 0.003 | -0.000 |
| | | | | (0.005) | (0.004) | (0.003) |
| D_trifecta | | | | 0.004 | 0.004 | 0.001 |
| | | | | (0.004) | (0.003) | (0.002) |
| ctz_ideo | | | | 0.000* | 0.000* | -0.000 |
| | | | | (0.000) | (0.000) | (0.000) |
| gvt_ideo | | | | 0.000 | 0.000 | 0.000 |
| | | | | (0.000) | (0.000) | (0.000) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | 0.624 | 0.784 | 0.630 | 0.638 | 0.647 | 0.784 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.2: Estimates of effect on log of union coverage

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | -0.164*** | -0.098* | -0.178*** | -0.154** | -0.163** | -0.150*** |
| | (0.051) | (0.052) | (0.054) | (0.057) | (0.059) | (0.050) |
| lgdp | | | 0.261* | | 0.281** | 0.232* |
| | | | (0.147) | | (0.135) | (0.128) |
| Dgdp | | | 0.082 | | -0.049 | -0.028 |
| | | | (0.197) | | (0.217) | (0.201) |
| emp | | | 1.020* | | 1.652*** | 1.027* |
| | | | (0.545) | | (0.447) | (0.508) |
| unemp | | | -1.005 | | -0.131 | -0.127 |
| | | | (1.019) | | (0.961) | (0.808) |
| manuf | | | -1.124 | | -0.267 | -0.366 |
| | | | (0.785) | | (0.665) | (0.659) |
| college | | | -1.043* | | -0.008 | -0.538 |
| | | | (0.605) | | (0.493) | (0.333) |
| house_dem | | | | 0.004 | 0.031 | -0.243*** |
| | | | | (0.142) | (0.127) | (0.088) |
| dgov | | | | -0.076 | -0.066* | -0.048 |
| | | | | (0.046) | (0.038) | (0.033) |
| R_legctrl | | | | -0.010 | -0.013 | 0.021 |
| | | | | (0.034) | (0.028) | (0.019) |
| D_legctrl | | | | -0.042 | -0.033 | -0.005 |
| | | | | (0.032) | (0.026) | (0.021) |
| R_trifecta | | | | 0.027 | 0.047 | 0.006 |
| | | | | (0.044) | (0.040) | (0.029) |
| D_trifecta | | | | 0.023 | 0.016 | 0.017 |
| | | | | (0.037) | (0.029) | (0.020) |
| ctz_ideo | | | | 0.001 | 0.001 | 0.000 |
| | | | | (0.001) | (0.001) | (0.001) |
| gvt_ideo | | | | 0.004*** | 0.004** | 0.002 |
| | | | | (0.001) | (0.001) | (0.001) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | 0.607 | 0.729 | 0.602 | 0.616 | 0.613 | 0.717 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: Estimates of effect on nonmember share

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | 0.030 | 0.019 | 0.026 | 0.021 | 0.017 | 0.015 |
| | (0.018) | (0.014) | (0.017) | (0.014) | (0.013) | (0.015) |
| lgdp | | | 0.012 | | -0.017 | 0.003 |
| | | | (0.062) | | (0.056) | (0.074) |
| Dgdp | | | 0.008 | | 0.024 | 0.025 |
| | | | (0.061) | | (0.062) | (0.056) |
| emp | | | -0.048 | | -0.007 | 0.051 |
| | | | (0.161) | | (0.179) | (0.148) |
| unemp | | | -0.429 | | -0.363 | -0.311 |
| | | | (0.256) | | (0.247) | (0.223) |
| manuf | | | -0.217 | | -0.185 | -0.093 |
| | | | (0.178) | | (0.189) | (0.202) |
| college | | | -0.159 | | -0.004 | -0.003 |
| | | | (0.098) | | (0.144) | (0.198) |
| house_dem | | | | -0.019 | -0.027 | 0.012 |
| | | | | (0.034) | (0.033) | (0.029) |
| dgov | | | | 0.007 | 0.007 | 0.006 |
| | | | | (0.010) | (0.011) | (0.011) |
| R_legctrl | | | | 0.007 | 0.007 | 0.002 |
| | | | | (0.009) | (0.008) | (0.008) |
| D_legctrl | | | | -0.022** | -0.021** | -0.012 |
| | | | | (0.009) | (0.010) | (0.010) |
| R_trifecta | | | | -0.013* | -0.011 | -0.005 |
| | | | | (0.007) | (0.008) | (0.008) |
| D_trifecta | | | | 0.018* | 0.016 | 0.013 |
| | | | | (0.009) | (0.010) | (0.011) |
| ctz_ideo | | | | 0.000 | 0.000 | -0.000 |
| | | | | (0.000) | (0.000) | (0.000) |
| gvt_ideo | | | | -0.000 | -0.000 | -0.000 |
| | | | | (0.000) | (0.000) | (0.000) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | 0.122 | 0.240 | 0.132 | 0.185 | 0.190 | 0.272 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.4: Estimates of effect on inflow of unionized workers

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | -0.049 | 0.353** | 0.049 | 0.272 | 0.288 | 0.282* |
| | (0.270) | (0.151) | (0.232) | (0.247) | (0.209) | (0.166) |
| lgdp | | | 2.084 | | 2.435 | 0.618 |
| | | | (1.486) | | (1.493) | (1.053) |
| Dgdp | | | -1.931 | | -2.456 | -2.231 |
| | | | (1.437) | | (1.746) | (1.430) |
| emp | | | -6.668 | | -7.665 | -0.698 |
| | | | (5.024) | | (5.702) | (2.875) |
| unemp | | | 1.591 | | 0.234 | -1.071 |
| | | | (6.798) | | (5.942) | (4.604) |
| manuf | | | -8.436 | | -6.821 | 1.615 |
| | | | (7.651) | | (7.425) | (5.601) |
| college | | | 8.476* | | 7.326 | 0.570 |
| | | | (4.306) | | (4.686) | (3.642) |
| house_dem | | | | 0.272 | 0.455 | -0.253 |
| | | | | (0.861) | (0.864) | (0.944) |
| dgov | | | | -0.099 | 0.061 | -0.047 |
| | | | | (0.321) | (0.338) | (0.277) |
| R_legctrl | | | | 0.154 | 0.119 | -0.009 |
| | | | | (0.338) | (0.297) | (0.166) |
| D_legctrl | | | | 0.182 | 0.135 | -0.003 |
| | | | | (0.207) | (0.192) | (0.146) |
| R_trifecta | | | | -0.156 | -0.157 | 0.002 |
| | | | | (0.292) | (0.264) | (0.257) |
| D_trifecta | | | | 0.183 | 0.170 | 0.285 |
| | | | | (0.266) | (0.254) | (0.282) |
| ctz_ideo | | | | 0.009 | 0.011 | 0.005 |
| | | | | (0.016) | (0.014) | (0.010) |
| gvt_ideo | | | | 0.010 | 0.000 | -0.005 |
| | | | | (0.009) | (0.011) | (0.006) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | 0.455 | 0.586 | 0.465 | 0.493 | 0.501 | 0.613 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.5: Estimates of effect on outflow of unionized workers

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | -0.012 | -0.014 | -0.010 | -0.036* | -0.025 | -0.036** |
| | (0.008) | (0.015) | (0.009) | (0.019) | (0.020) | (0.016) |
| lgdp | | | 0.239** | | 0.195 | 0.318 |
| | | | (0.113) | | (0.115) | (0.254) |
| Dgdp | | | -0.198 | | -0.350* | -0.294 |
| | | | (0.212) | | (0.197) | (0.234) |
| emp | | | 0.108 | | 0.131 | -0.022 |
| | | | (0.314) | | (0.292) | (0.310) |
| unemp | | | 0.159 | | -0.019 | -0.135 |
| | | | (0.542) | | (0.528) | (0.563) |
| manuf | | | -0.230 | | -0.527 | -0.418 |
| | | | (0.562) | | (0.601) | (0.692) |
| college | | | -0.173 | | -0.337 | -0.374 |
| | | | (0.212) | | (0.379) | (0.362) |
| house_dem | | | | -0.116 | -0.108 | -0.269* |
| | | | | (0.104) | (0.117) | (0.144) |
| dgov | | | | 0.006 | 0.003 | 0.016 |
| | | | | (0.022) | (0.024) | (0.029) |
| R_legctrl | | | | -0.044*** | -0.045*** | -0.026 |
| | | | | (0.014) | (0.016) | (0.020) |
| D_legctrl | | | | 0.012 | 0.017 | 0.029 |
| | | | | (0.016) | (0.017) | (0.027) |
| R_trifecta | | | | 0.049* | 0.054 | 0.035 |
| | | | | (0.028) | (0.037) | (0.042) |
| D_trifecta | | | | -0.040** | -0.047** | -0.055** |
| | | | | (0.019) | (0.020) | (0.026) |
| ctz_ideo | | | | -0.000 | -0.000 | -0.002 |
| | | | | (0.001) | (0.001) | (0.001) |
| gvt_ideo | | | | 0.000 | 0.001 | 0.000 |
| | | | | (0.001) | (0.002) | (0.002) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | 0.015 | 0.000 | 0.015 | 0.030 | 0.030 | 0.035 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.6: Estimates of effect on union wage differentials

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | 0.020 | 0.025** | 0.022* | 0.023 | 0.021* | 0.036*** |
| | (0.017) | (0.012) | (0.012) | (0.014) | (0.012) | (0.012) |
| lgdp | | | -0.081* | | -0.076* | -0.089 |
| | | | (0.047) | | (0.043) | (0.061) |
| Dgdp | | | -0.028 | | -0.040 | -0.056 |
| | | | (0.099) | | (0.109) | (0.112) |
| emp | | | 0.159 | | 0.288* | 0.321* |
| | | | (0.176) | | (0.167) | (0.174) |
| unemp | | | 0.170 | | 0.339 | 0.385 |
| | | | (0.292) | | (0.283) | (0.299) |
| manuf | | | -0.480*** | | -0.393** | -0.360 |
| | | | (0.166) | | (0.173) | (0.250) |
| college | | | -0.184 | | -0.072 | -0.091 |
| | | | (0.140) | | (0.125) | (0.158) |
| house_dem | | | | 0.049 | 0.027 | 0.021 |
| | | | | (0.046) | (0.047) | (0.055) |
| dgov | | | | -0.009 | -0.012 | -0.011 |
| | | | | (0.013) | (0.012) | (0.013) |
| R_legctrl | | | | 0.004 | 0.002 | -0.000 |
| | | | | (0.013) | (0.013) | (0.012) |
| D_legctrl | | | | 0.002 | -0.000 | -0.004 |
| | | | | (0.008) | (0.008) | (0.008) |
| R_trifecta | | | | -0.003 | -0.001 | -0.002 |
| | | | | (0.013) | (0.012) | (0.013) |
| D_trifecta | | | | 0.005 | 0.005 | -0.000 |
| | | | | (0.009) | (0.009) | (0.009) |
| ctz_ideo | | | | -0.001* | -0.001** | -0.001** |
| | | | | (0.000) | (0.000) | (0.000) |
| gvt_ideo | | | | 0.000 | 0.000 | 0.000 |
| | | | | (0.000) | (0.000) | (0.001) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | 0.104 | 0.166 | 0.124 | 0.100 | 0.113 | 0.155 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.7: Estimates of effect on non-union wages

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | -0.034** | -0.009 | -0.024** | -0.036** | -0.024** | -0.004 |
| | (0.014) | (0.010) | (0.010) | (0.013) | (0.010) | (0.011) |
| lgdp | | | 0.318*** | | 0.300*** | 0.282*** |
| | | | (0.063) | | (0.065) | (0.064) |
| Dgdp | | | -0.213*** | | -0.231*** | -0.234*** |
| | | | (0.064) | | (0.057) | (0.058) |
| emp | | | -0.359** | | -0.445*** | -0.348*** |
| | | | (0.143) | | (0.126) | (0.100) |
| unemp | | | -0.071 | | -0.199 | 0.109 |
| | | | (0.212) | | (0.179) | (0.166) |
| manuf | | | 0.511** | | 0.458** | -0.110 |
| | | | (0.187) | | (0.181) | (0.159) |
| college | | | 0.897*** | | 0.889*** | 0.993*** |
| | | | (0.112) | | (0.120) | (0.137) |
| house_dem | | | | -0.045 | -0.006 | -0.047 |
| | | | | (0.044) | (0.028) | (0.032) |
| dgov | | | | -0.022* | -0.018 | -0.010 |
| | | | | (0.013) | (0.012) | (0.010) |
| R_legctrl | | | | 0.014 | 0.015* | 0.010* |
| | | | | (0.011) | (0.009) | (0.006) |
| D_legctrl | | | | -0.013 | -0.016 | -0.008 |
| | | | | (0.010) | (0.010) | (0.009) |
| R_trifecta | | | | -0.019 | -0.023* | -0.007 |
| | | | | (0.014) | (0.013) | (0.008) |
| D_trifecta | | | | 0.014 | 0.014 | 0.003 |
| | | | | (0.012) | (0.011) | (0.009) |
| ctz_ideo | | | | 0.000 | 0.000** | 0.001** |
| | | | | (0.000) | (0.000) | (0.000) |
| gvt_ideo | | | | 0.001 | 0.000 | 0.000 |
| | | | | (0.000) | (0.000) | (0.000) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | 0.974 | 0.983 | 0.982 | 0.978 | 0.983 | 0.987 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.8: Estimates of effect on union wages

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | -0.025 | 0.018 | -0.017 | -0.015 | -0.008 | 0.032 |
| | (0.020) | (0.025) | (0.021) | (0.024) | (0.024) | (0.024) |
| lgdp | | | 0.090 | | 0.097 | 0.047 |
| | | | (0.100) | | (0.105) | (0.099) |
| Dgdp | | | -0.049 | | -0.072 | -0.064 |
| | | | (0.139) | | (0.149) | (0.143) |
| emp | | | -0.228 | | -0.143 | -0.122 |
| | | | (0.225) | | (0.270) | (0.246) |
| unemp | | | -0.235 | | -0.042 | 0.182 |
| | | | (0.318) | | (0.340) | (0.291) |
| manuf | | | -0.208 | | -0.096 | -0.200 |
| | | | (0.263) | | (0.264) | (0.291) |
| college | | | $0.494^{***}$ | | $0.556^{**}$ | $0.579^{***}$ |
| | | | (0.157) | | (0.205) | (0.186) |
| house_dem | | | | 0.067 | 0.079 | 0.036 |
| | | | | (0.064) | (0.057) | (0.072) |
| dgov | | | | -0.026 | -0.026 | -0.021 |
| | | | | (0.020) | (0.019) | (0.018) |
| R_legctrl | | | | 0.009 | 0.009 | 0.004 |
| | | | | (0.012) | (0.012) | (0.013) |
| D_legctrl | | | | $-0.027^{*}$ | $-0.031^{**}$ | $-0.029^{*}$ |
| | | | | (0.015) | (0.014) | (0.017) |
| R_trifecta | | | | -0.013 | -0.014 | -0.010 |
| | | | | (0.015) | (0.013) | (0.016) |
| D_trifecta | | | | $0.037^{**}$ | $0.037^{***}$ | $0.027^{*}$ |
| | | | | (0.014) | (0.012) | (0.015) |
| ctz_ideo | | | | -0.000 | -0.000 | -0.001 |
| | | | | (0.001) | (0.000) | (0.001) |
| gvt_ideo | | | | 0.000 | 0.000 | 0.000 |
| | | | | (0.001) | (0.001) | (0.001) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | 0.932 | 0.938 | 0.925 | 0.933 | 0.927 | 0.933 |

Standard errors in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

Table B.9: Estimates of effect on strike participation

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | -0.157 | 5.434 | 2.656 | 2.505 | 4.977 | 3.286 |
| | (3.698) | (4.701) | (3.230) | (4.839) | (4.656) | (4.690) |
| lgdp | | | -6.227 | | -5.759 | -23.453 |
| | | | (11.209) | | (13.482) | (18.461) |
| Dgdp | | | -46.852 | | -23.589 | -23.271 |
| | | | (30.182) | | (21.148) | (21.530) |
| emp | | | -87.792 | | -112.062 | -102.438 |
| | | | (63.745) | | (71.445) | (77.038) |
| unemp | | | -265.623** | | -283.326* | -285.151* |
| | | | (128.937) | | (149.107) | (151.337) |
| manuf | | | 34.828 | | 20.887 | -60.309 |
| | | | (60.896) | | (72.150) | (78.240) |
| college | | | 71.837* | | 21.433 | 4.453 |
| | | | (41.620) | | (74.515) | (71.006) |
| house_dem | | | | 0.874 | 3.522 | -5.969 |
| | | | | (6.998) | (7.429) | (9.511) |
| dgov | | | | -9.983** | -9.852** | -10.820** |
| | | | | (4.781) | (4.202) | (4.645) |
| R_legctrl | | | | 3.178 | 3.660 | 1.789 |
| | | | | (2.668) | (2.704) | (3.007) |
| D_legctrl | | | | -5.665 | -5.208 | -7.588* |
| | | | | (3.450) | (3.678) | (4.084) |
| R_trifecta | | | | -7.131 | -6.815 | -7.449 |
| | | | | (4.820) | (4.493) | (5.285) |
| D_trifecta | | | | 4.906 | 3.897 | 5.420 |
| | | | | (3.865) | (3.517) | (4.355) |
| ctz_ideo | | | | -0.211 | -0.169 | -0.219 |
| | | | | (0.147) | (0.154) | (0.158) |
| gvt_ideo | | | | 0.386** | 0.409** | 0.426** |
| | | | | (0.170) | (0.172) | (0.180) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | 0.039 | 0.059 | 0.048 | 0.051 | 0.055 | 0.071 |

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.10: Estimates of effect on strike duration

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| RTW | -3.477 | -9.487 | -3.784 | -1.483 | -1.680 | -2.533 |
| | (4.971) | (7.544) | (5.304) | (6.780) | (7.679) | (10.557) |
| lgdp | | | -11.328 | | -15.694 | -67.236* |
| | | | (28.755) | | (30.027) | (33.688) |
| Dgdp | | | -164.485** | | -168.024** | -121.740 |
| | | | (62.734) | | (68.320) | (78.753) |
| emp | | | 13.992 | | -26.037 | -44.652 |
| | | | (116.626) | | (127.121) | (146.881) |
| unemp | | | 109.642 | | 76.614 | -92.665 |
| | | | (171.782) | | (194.630) | (213.744) |
| manuf | | | 15.254 | | 42.481 | 191.901 |
| | | | (131.619) | | (140.831) | (146.291) |
| college | | | 42.714 | | 28.903 | -36.238 |
| | | | (86.368) | | (122.270) | (172.190) |
| house_dem | | | | 12.824 | 4.311 | 38.967 |
| | | | | (23.074) | (23.371) | (33.187) |
| dgov | | | | 6.436 | 4.420 | 1.281 |
| | | | | (6.691) | (6.954) | (9.144) |
| R_legctrl | | | | 6.869 | 3.520 | 1.243 |
| | | | | (6.838) | (6.446) | (7.178) |
| D_legctrl | | | | 2.864 | 3.322 | 3.944 |
| | | | | (5.935) | (6.787) | (6.109) |
| R_trifecta | | | | -5.832 | -4.107 | -0.899 |
| | | | | (6.668) | (6.809) | (7.717) |
| D_trifecta | | | | -7.465 | -6.600 | -5.661 |
| | | | | (6.396) | (7.007) | (7.758) |
| ctz_ideo | | | | -0.106 | -0.200 | -0.787** |
| | | | | (0.246) | (0.256) | (0.372) |
| gvt_ideo | | | | -0.028 | 0.053 | 0.267 |
| | | | | (0.365) | (0.410) | (0.482) |
| Time trends | No | Yes | No | No | No | Yes |
| No. of states | 30 | 30 | 30 | 29 | 29 | 29 |
| Adjusted $R^2$ | -0.008 | 0.008 | 0.001 | -0.008 | -0.003 | 0.019 |

Standard errors in parentheses

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$