University of Pennsylvania
ScholarlyCommons

Publicly Accessible Penn Dissertations

1-1-2016

Essays in Political Economy

Devin Joseph Reilly
University of Pennsylvania, reillyde@sas.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/edissertations

Part of the Economics Commons

Essays in Political Economy

Abstract

This dissertation consists of two chapters on topics in political economy. In the first chapter, using hand-collected data from collective bargaining agreements for state-level public sector unions, I develop and calibrate a stochastic bargaining model as in Merlo and Wilson (1995) and investigate the effects of political and economic variables on wages, pensions, and delay in reaching agreement. In the model, a governor and a union bargain over a wage and pension outcome. The economic state, as measured by the unemployment rate, evolves stochastically and affects the propensity of the governor and union to reach agreement in any given period. Furthermore, political variables, including party of the governor, partisanship of the district, and incumbency, affect the relative payoffs and therefore the wage, pension, and time to agreement. I find that negotiated wage and pension growth is higher under Democratic governors, while increases in the unemployment rate at the beginning of bargaining have a negative impact on compensation levels, the magnitude of which varies by party and time before the next election.

In the second chapter, which is co-authored with Ekim Cem Muyan, I develop a model of campaign strategies, namely the choice to campaign negatively or positively. In particular, I construct a model of political campaigns, based off of Skaperdas and Grofman (1995), in which candidates allocate their budget between positive and negative campaigning. Elections vary according to politician- and district-specific characteristics, as well as the unobservable (to the econometrician) measure of voter types. I calibrate the model to match stylized facts on campaign tone that we document using a wide array of sources, including data on advertising tone from Wisconsin Advertising Project, campaign contributions from the Database on Ideology, Money in Politics, and Elections, and election results. The calibrated model implies that, overall, campaign spending is not particularly effective at increasing votes -- a 10% increase in the average candidate's budget, corresponding to about $240,000, raises his or her expected vote share by about 0.4 percentage points. The model also implies that negativity is marginally more useful for candidates who are trailing than those leading, though not by a wide margin.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Economics

First Advisor
Holger Sieg

Keywords
Political economy, Public economics

Subject Categories
Economics

ESSAYS IN POLITICAL ECONOMY

Devin Reilly

A DISSERTATION

in

Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

_____

Holger Sieg, Professor of Economics

Graduate Group Chairperson

_____

Jesus Fernandez-Villaverde, Professor of Economics

Dissertation Committee

Camilo Garcia-Jimeno, Professor of Economics
Antonio Merlo, Professor of Economics
Holger Sieg, Professor of Economics

To my family and Jessica for their unwavering support.

ACKNOWLEDGEMENTS

ABSTRACT

ESSAYS IN POLITICAL ECONOMY

Devin Reilly

Holger Sieg

This dissertation consists of two chapters on topics in political economy. In the first chapter, using hand-collected data from collective bargaining agreements for state-level public sector unions, I develop and calibrate a stochastic bargaining model as in Merlo and Wilson (1995) and investigate the effects of political and economic variables on wages, pensions, and delay in reaching agreement. In the model, a governor and a union bargain over a wage and pension outcome. The economic state, as measured by the unemployment rate, evolves stochastically and affects the propensity of the governor and union to reach agreement in any given period. Furthermore, political variables, including party of the governor, partisanship of the district, and incumbency, affect the relative payoffs and therefore the wage, pension, and time to agreement. I find that negotiated wage and pension growth is higher under Democratic governors, while increases in the unemployment rate at the beginning of bargaining have a negative impact on compensation levels, the magnitude of which varies by party and time before the next election.

In the second chapter, which is co-authored with Ekim Cem Muyan, I develop a model of campaign strategies, namely the choice to campaign negatively or positively. In particular, I construct a model of political campaigns, based off of Skaperdas and Grofman (1995), in which candidates allocate their budget between positive and negative campaigning. Elections vary according to politician- and district-specific characteristics, as well as the unobservable (to the econometrician) measure of voter types. I calibrate the model to match stylized facts on campaign tone that we document using a wide array of sources,

including data on advertising tone from Wisconsin Advertising Project, campaign contributions from the Database on Ideology, Money in Politics, and Elections, and election results. The calibrated model implies that, overall, campaign spending is not particularly effective at increasing votes – a 10% increase in the average candidate's budget, corresponding to about \$240,000, raises his or her expected vote share by about 0.4 percentage points. The model also implies that negativity is marginally more useful for candidates who are trailing than those leading, though not by a wide margin.

Contents

List of Tables

List of Figures

Chapter 1 : The Political Economy of Public Sector Union Bargaining

1.1 Introduction

In the United States, state and local governments spent about $2.4 trillion in 2014, which constituted 13.7% of GDP. Employee compensation accounted for over half of this, totaling over $1.3 trillion.[1] Furthermore, the level, growth, and structure of public sector compensation are all subject to factors distinct from private sector labor markets.[2] The scale of public sector compensation alone suggests that understanding its determinants is important. Moreover, the recent financial crisis has strained state budgets significantly, with tax revenues declining about 12% in real terms in 2009 and only recently reaching their pre-recession levels. Additionally, recent work suggests that state pension plans, which cover over 22 million participants and hold over $2.4 trillion in assets, are significantly underfunded and upcoming pension obligations pose concerns for state finances in the near

---

[1] See U.S. Bureau of Economic Analysis, National Income and Product Accounts, Tables 3.3 and 6.2D.

[2] Certainly since there is mobility between public and private sector markets, the pressures are not completely distinct. However, several studies, including for instance Bewerunge and Rosen (2013), have documented how the structure and levels of public sector compensation are different from the private sector controlling for various worker observables. Additionally, Bagchi (2015) has documented a positive relationship between election closeness and pension generosity for municipal workers in Pennsylvania, suggesting further that political variables may play a role in determining public sector compensation.

future.[3] Thus, the question of what determines public sector compensation has only become more important in recent years.

While the topic of public sector compensation is not new, most previous work has focused on either measuring the difference between public and private sector compensation, or on the effect of union strength on either pensions or wages.[4] However, relatively little is known about the political factors that affect public sector compensation, especially in the context of collective bargaining.[5] Anecdotal evidence suggests that political factors may be important in determining bargaining outcomes. For instance, there is an increasing partisan divide over public sector compensation, with Republican governors such as Scott Walker and Chris Christie taking harder lines with unions than many Democratic governors.[6] Furthermore, the few papers that focus on politics and public sector union outcomes show that political factors may be important. For instance, Sieg and Wang (2013) focus on mayoral elections and find that union endorsements can affect election outcomes, and therefore that such endorsements can be used to extract more union-friendly fiscal policies. Additionally, Bagchi (2015) finds that election closeness is positively related to pension generosity using data from municipalities in Pennsylvania, while Anzia and Moe (2015a) find differences in pension expansions and reductions across political parties. Thus, further investigation of the interaction of politics and public sector union bargaining outcomes is needed.

Using a novel dataset with information on collective bargaining agreements for state public sector unions, this chapter investigates the effects of political and economic variables on bargaining outcomes. In particular, I develop a stochastic bargaining model as in Merlo and Wilson (1995) and Merlo (1997) in which a governor and a union bargain over a wage and pension outcome. The economic state, as measured by the unemployment rate, evolves stochastically and affects the propensity of the governor and union to reach agreement in any

---

[3]See, for instance, Novy-Marx and Rauh (2009), (2011), and (2014).

[4]For pensions, see for instance Munnell and Quinby (2011a). For wages, see, for example, Freeman and Han (2012) and Anzia and Moe (2015b).

[5]See Sieg and Wang (2013), Bagchi (2015), and Anzia and Moe (2015a) as exceptions, which I discuss below.

[6]See *New York Times Magazine*, June 12, 2015, "Scott Walker and the Fate of the Union."

given period. Furthermore, political variables, including party of the governor, partisanship of the district, and incumbency affect the relative payoffs and therefore the wage, pension, and delay in reaching agreement.

In particular, each period the unemployment rate is realized and a proposer, either the governor or the union, is randomly selected. The proposer makes a wage and pension offer to the responder, who accepts or rejects. All else equal, the union prefers to receive larger wages and pensions, while the governor prefers lower wages and pensions. The relative costs to the governor from raising wages and pensions depend on the economic state and his political party. Furthermore, both parties receive a flow benefit during bargaining. The union's benefit depends on the current wages and pension. The governor's depends on whether or not the parties have reached agreement, with the level of benefits depending on political party and economic state. For instance, Democrats may have relatively high benefits from reaching agreement especially in good economic states, whereas Republicans on net may be relatively better off "bargaining tough" with unions, especially when the unemployment rate is high. Finally, if the sides reach the nearest election date without agreement, both sides may face a disagreement cost.

In some states and dates, parties may mutually benefit from postponing agreement until a better state arises as raising compensation when the unemployment rate is high may be particularly costly. The propensity to delay, as well as the realized wage and pension agreements, depend on the relative values of flow benefits, costs and benefits of raising compensation, and terminal disagreement costs. Thus, model outcomes will depend on political party, proximity to the next election, and economic state, among other variables.

The logic of the model is as follows. Politicians may be more reluctant to increase employee compensation in bad economic times. This could be due in part to voter pressure and in part due to budgetary effects. Thus, if bargaining begins in a bad economic state, politicians and unions may both be willing to postpone reaching agreement until a later period in which there is less political pressure. The framework of a stochastic bargaining model is known

to explain delay in reaching agreement, which is prevalent in the data, and can capture these mechanisms, shedding light on several factors that determine public sector bargaining outcomes.

I find that overall compensation growth is larger under Democrats and in better economic states, as measured by the unemployment rate. In particular, I find that wage growth under Democrats is approximately 0.43 percentage points larger than than under Republicans, while pension growth is about 1.33 percentage points larger. Delay in reaching agreement is also significantly lower under Democrats by approximately 1.45 months. Additionally, the model implies that increases in the unemployment rate at the time bargaining begins tend to decrease both wages and pensions. Simulations from the calibrated model show that, for a representative bargaining spell, an approximately one percentage point decrease in initial unemployment rate generates higher wage growth by about 0.53 percentage points for Democrats or 0.44 percentage points for Republicans. Pension growth responds by about 0.31 percentage points for Democrats and 0.06 percentage points for Republicans, depending on the level of unemployment. Finally, as consistent with previous work, pension growth is larger for bargaining spells that begin with more well-funded pension plans.

This chapter contributes to three main strands of the literature. First, the model is a version of the general stochastic bargaining model developed in Merlo and Wilson (1995), which can rationalize delay in bargaining settings. Most applications of the model, including Merlo (1997) and Diaz-Moreno and Galdon-Sanchez (2005), assume transferable utility and bargaining over a unidimensional, perfectly divisible object.[7] In this chapter, however, the governor and union have different, non-linear utilities over bargaining outcomes, which are multi-dimensional wage-pension pairs. Non-transferable utility makes sense in this case relative to other applications, since the union cares about the wage and pensions insofar as they increase consumption, while the governor cares about wages and pensions primarily through its effect on elections and the state budget. Thus, whereas in private union bar-

---

[7]See Diaz-Moreno and Galdon-Sanchez (2000) for a counterexample, which features non-transferable utility. In particular, the though players have linear utility, the model allows for players to have different discount factors.

gaining the object is profits, which both sides may value more or less equally, in the public sector the utilities of the two sides likely differ. This chapter is also, to my knowledge, the first application of stochastic bargaining models to *public*-sector union bargaining.

Second, the chapter contributes to a wide literature on collective bargaining. Much of the older literature on collective bargaining looks at the private sector and the prevalence of strikes and holdouts. For example, Farber (1978), Hart (1989), and Kennan and Wilson (1989) focus on strike, whereas Gu and Kuhn (1998) explain holdouts. These papers often involve asymmetric information, either in regards to the cost of delay or firm profitability. Here, the model features perfect and complete information, rationalizing delay via a stochastic state that evolves over time. If the state is bad, it may make sense for sides to postpone and reach agreement later. Asymmetric information may be less suitable in explaining holdouts and delay in the public sector, since there is no "firm profitability" involved, and since information is often publicly available to both sides unlike in the private sector. Additionally, much of the more recent literature investigates the effect of unionization or collective bargaining laws on compensation in the public sector (see Falch and Strom (2005), Munnell and Quinby (2011a), Feiveson (2015), and Anzia and Moe (2015b) for instance). While this chapter does not directly compare compensation across different collective bargaining regimes, it does more explicitly model the bargaining process, which sheds light on how compensation is determined within states that allow bargaining.

Finally, this chapter contributes to the literature on the interaction between politics and public sector unions and compensation. Sieg and Wang (2013), as previously discussed, focuses on the effect of union endorsements on mayoral election outcomes, as well as how endorsements can be used to extract concessions in the form of pro-union policy. Bagchi (2015) analyzes municipalities in Pennsylvania and the effect of political factors on pension funding, finding higher underfunding in politically competitive jurisdictions. It also provides evidence that pension generosity is higher in more competitive municipalities. Finally, Anzia and Moe (2015a) analyze pension expansions and restrictions and finds differences between Republican and Democratic support for such changes. In contrast to these three,

this chapter focuses explicitly on the bargaining process between unions and governments. Furthermore, I analyze the joint determination of wages, pensions, and bargaining delay, whereas these papers look at either pensions or, in the case of Sieg and Wang (2013), total payrolls.

Section 1.2 discuss some of the background and institutional details of public sector union bargaining. Section 1.3 discusses the data used and some evidence on the relationship between public sector union bargaining outcomes and political and economic variables. Section 1.4 describes the model, while Section 1.5 discusses the calibration procedure and fit. Section 1.6 further decomposes the model implications, and section 1.7 concludes the main text of the chapter. Finally, section 1.8 provides the appendix.

1.2   Institutional Details

In this section, I briefly discuss some of the relevant institutional details, namely collective bargaining and public sector unions. In the private sector, labor unions have diminished in importance over the past several decades, with membership decreasing from 24.2% of the workforce in 1973 to 6.6% in 2014. However, in the public sector, membership has increased from 23.0% to 35.7% over the same period.[8] The rate is even higher for public safety workers. The unionization rate among police officers is 60%, and among firefighters it is 67%.[9] Furthermore, almost 40% of all public employees are covered by a collective bargaining agreement. Thus, collective bargaining plays an important role in determining compensation and other terms of employment in the public sector.

Collective bargaining and wage negotiation is allowed in 40 states for general employees, with some slight differences for teachers or public safety employees.[10] However, the vast majority of states do not allow employees to strike. For public safety employees, only two

---

[8]See Hirsch and Macpherson (2014).
[9]See Sanes and Schmitt (2014).
[10]See Sanes and Schmitt (2014).

states – Hawaii and Ohio – allow striking, while for general employees only 10 do. Even in states in which striking is allowed, it is incredibly rare, particularly at the state level.[11]

In collective bargaining in the public sector, the union and the state representative bargain over various terms of employment, often including wages, pension benefits, health benefits, employment, and many others. The agreements can be up to 200 pages or more, though many of these can be filled with minor details and legalese. The union represents a certain set of employees within the public sector and there are multiple public unions within each state, often with public safety employees (i.e. police and firefighters), teachers, and other employees grouped together. Within a given union, the specified lengths of the contracts are typically the same over time. For instance, for most unions in New Jersey, contracts are four years long and there is very little deviation from this. Thus, contract length can be seen as exogenous.

Each union at any given point in time will be under a particular contract. Prior to its expiration, the collective bargaining process begins. If a new contract is not agreed upon prior to the expiration date of the old contract, then the old contract remains in place until agreement is reached. As I discuss in section 1.3, while many bargaining spells reach agreement prior to expiration date, the majority do not. Once the agreement is reached, the new contract begins and bargaining stops until close to the following expiration date. Note that the governor of the state is required to sign all agreements, and thus is a party to all negotiations. Thus, in the model, I focus only on the governor on the side of the state.[12]

Finally, an important point regarding compensation is the structure of pension plans. Pension in the private sector are generally defined-contribution. In such plans, the agreements between workers and firms specify an amount that the firm will contribute to a retirement fund. Then, the worker can choose how to allocate the funds in a portfolio and is entitled

---

[11]See, for example, *Governing Magazine*, October 10, 2012, "Why Public-Sector Strikes Are So Rare."

[12]In some states, the legislature has a vote over contracts, and in many states they vote on pension changes. However, for tractability, and because the governor has veto power and thus likely exerts a disproportionate amount of power on the negotiations, I focus exclusively on the governor.

to the principal and returns, but is not guaranteed any set amount in retirement – hence defined-contribution. In the public sector, the majority of plans are defined-benefit.[13] In these plans, a formula is specified that guarantees a set benefit in retirement as a function of worker-specific variables, tenure and average salary, and plan-specific variables. I describe more precisely the formulation of these plans in the next section. Public workers and the government typically contribute a percentage of salary to a pension fund that invests as it sees prudent.[14] However, unlike defined-contribution plans, in defined-benefit plans the worker is entitled to whatever the specified *benefit* is, not to the returns of the pension fund. Since plans can potentially be underfunded and benefits do not need to be paid out until the future when a worker retires, there may be incentives for the government to provide generous pensions, consistent with previous research.[15]

1.3   Data

The primary sources of data are state-level collective bargaining (CB) agreements for various years and bargaining units. From each contract, I hand-collect relevant information including salary increases, agreement date, coverage dates, and other terms of employment. Some contracts also contain data on pension plans, though other times this is only available through state websites. For each state, contract availability does vary, although overall there is a bias towards more recent agreements.

For empirical implementation, the unit of analysis is a bargaining spell, which begins at the contract begin date and ends at either: (i) the date of agreement, or (ii) the date of the subsequent gubernatorial election, whichever is earlier. Those that do not reach agreement by the election are referred to as "ending in disagreement." The final sample contains 422

---

[13]Both in the aggregate and among the states in my sample, defined-benefit plans are the dominant form of pensions.
[14]The funding of plans by the government has historically been too low relative to what many experts believe is appropriate. See Novy-Marx and Rauh (2009), among many others, for a discussion on pension accounting and funding.
[15]See, for instance, Bagchi (2015).

bargaining spells. Note that if an observation ends in disagreement, there will be another observation for the same contract starting after the election and ending in agreement.[16] In the final sample, 39 bargaining spells end in disagreement, while in 383 spells the parties reach agreement. The final sample includes those spells remaining after I drop contracts in which there was no indication of the agreement date or the wage change, or those with begin dates earlier than 2000.

To supplement information from the CB contracts, I use pension data from the Public Plans Database, which includes data on funding levels for 107 state pension funds (which account for 90% of all state government pension assets and members) from 2001 to 2013. I also use state-level data on unionization rates,[17] demographic data from the Census and CPS, monthly unemployment rates from FRED, and electoral and political data. The electoral and political variables include gubernatorial party indicators, candidates, election outcomes, and the Cook Partisan Voting Index (CPVI), which is a measure of the partisanship of each state.[18]

1.3.1 Wages and Delay Times

Each CB agreement contains information on wage increases for covered employees. Generally, wage increases are the same for all workers. As an example, a contract covering a given union for a fixed time period specifies a nominal percent salary increase at particular months. Figure 1 shows an excerpt from the contract of the judiciary support staff unit in New Jersey for 2004-2008.

Figure 2 presents the distribution of gross average real wage growth in the sample.[19] These

---

[16]Technically such a spell could again end in disagreement again at the following election date (which would be four years away). However, none of the bargaining spells in the sample last this long.

[17]See Hirsch and Macpherson (2014).

[18]The CPVI is calculated as the average two-party vote share for the party that won the state in the previous two presidential elections minus the national average for that party. For instance, in 2004 and 2008, the national average two-party vote share was 51.2% to 49.8% in favor of Democrats. In Iowa, in 2004 the Democratic two-party vote share was 49.7%, while in 2008 it was 54.8%. Then, the average Democratic vote share over these two elections is 52.25%, making the CPVI Democrat+1.

[19]There are only 383 observations since, for the 39 bargaining spells ending in disagreement, there is no

Figure 1

EXAMPLE SALARY INCREASES, NJ JUDICIARY SUPPORT STAFF UNION, 2004-2008

A. Across the Board Salary Increases

7/1/04 (pay period 15)	2.0%
7/1/05 (pay period 15)	0.9%
7/1/06 (pay period 15)	2.0%
1/1/07 (pay period 2)	2.0%
7/1/07 (pay period 15)	2.25%
1/1/08 (pay period 2)	2.35%

are gross real wage increases, so an observation of 1.02 indicates a net real wage increase of about 2%. On average, real wages are essentially constant over the sample, with changes ranging from about a 5% decline to a 4.65% increase.

Figure 2

DISTRIBUTION OF REAL AVERAGE GROSS WAGE INCREASES, CONDITIONAL ON AGREEMENT

Distribution of Real Avg. Gross Wage Change
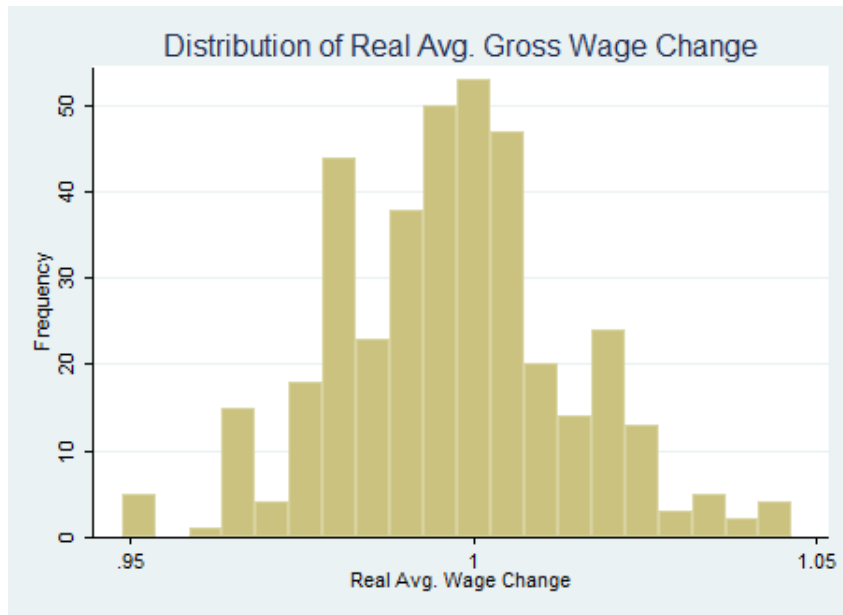
The contracts also show that bargaining spells are often characterized by significant delays in reaching agreement. Delay is defined as number of months between the end date of the previous contract and the agreement date of the new contract.[20] Figure 3 presents

---

agreed upon wage increase.

[20]Technically, since bargaining begins sometime before the old contract ends, the true "delay" is from the negotiation begin date to agreement date. However, the date at which the negotiation begins is not

10

the distribution of delay (in months). In the sample, average delay is approximately 6.91 months, while the maximum is 36 months. Note also the spike at 0 months, indicating that the modal contract is agreed to at or before the end of the previous contract.

Figure 3
DISTRIBUTION OF AGREEMENT DELAY TIMES
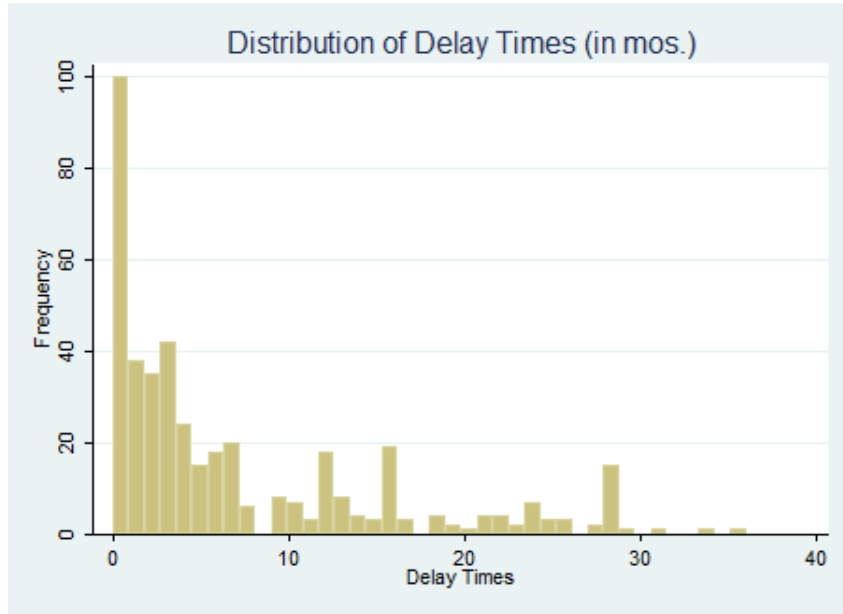
While there is a mass of bargaining spells with no delay (23.7% of spells have a delay of zero), most contracts have positive delay, implying that most contracts are finalized *after* the contract begins. This is certainly costly to the union in the sense of foregone wage increases, and is also likely costly to the politician. The median delay is only three months, but the maximum delay is 36 months, or three years. The prevalence of delay in public sector union bargaining in part motivates the decision to model the interaction as a stochastic bargaining game, which can rationalize delay in equilibrium.[21]

---

observable for all bargaining spells. Therefore, I measure delay as the time between the beginning of the contract until agreement. It is important to note that the begin date of the contract is also a logical date to start given the modeling framework, as delay prior to the beginning of the contract is essentially costless.

[21]See, for instance, Merlo and Wilson (1995) and Merlo (1997).

1.3.2 Pensions

Most public sector pensions are defined-benefit plans in which benefits are determined by a set formula. This formula is the product of three variables, some plan-specific and some employee-specific. The first is the retirement factor, which generally varies between 0.01 and 0.03. The second is the final average salary (FAS) of the worker. While this in part depends on the worker-specific salary path, the number of years included in the FAS is plan-specific, which I refer to as "FAS years." Note that in general, FAS years is negatively related to pension generosity, since most workers' salaries are increasing over their tenure.[22] The final variable is number of years of service, which is purely worker-specific. The product of these three variables will give the pre-tax monthly pension benefit for the worker. Furthermore, each plan specifies a contribution rate for the employee, which is the amount of pre-tax salary contributed by the employee during his working life. The contribution rate is often related implicitly to the pension generosity in that, on average, more generous pensions have higher contribution rates. However, there is no direct link as these are defined-benefit plans. Thus, an increase in the contribution rate will, all else equal, decrease the effective wage of the employee and not change the expected pension. Figures 4(a) to 4(c) show the distributions of each pension parameter.

To summarize pension generosity from these parameters, I compute "replacement rates" for each contract. In some contracts, pensions are not directly bargained over. For these spells, I use any changes during the span of the contract as the measure of pensions. The replacement rate is the after-tax pension income a worker takes home relative to the after-tax and after-contribution final year of salary for a specified worker. To measure this, I must assume some worker-specific characteristics. I consider an unmarried state public safety worker with 30 years of tenure and a $50,000 pre-tax retirement salary.[23] Figure

---

[22]Even beyond this, the FAS is usually specified to be the average of the, say, four *highest* years of salary, not last working years. So, even if a worker's salary declines as he nears retirement, the pension will usually be tied to the peak salary over his or her working life.

[23]The fact that the worker is unmarried does not affect the pre-tax pension benefit, but may affect the tax rate the worker is subject to. Appendix 1.8.3 describes the computation of pensions and tax rates in greater detail.
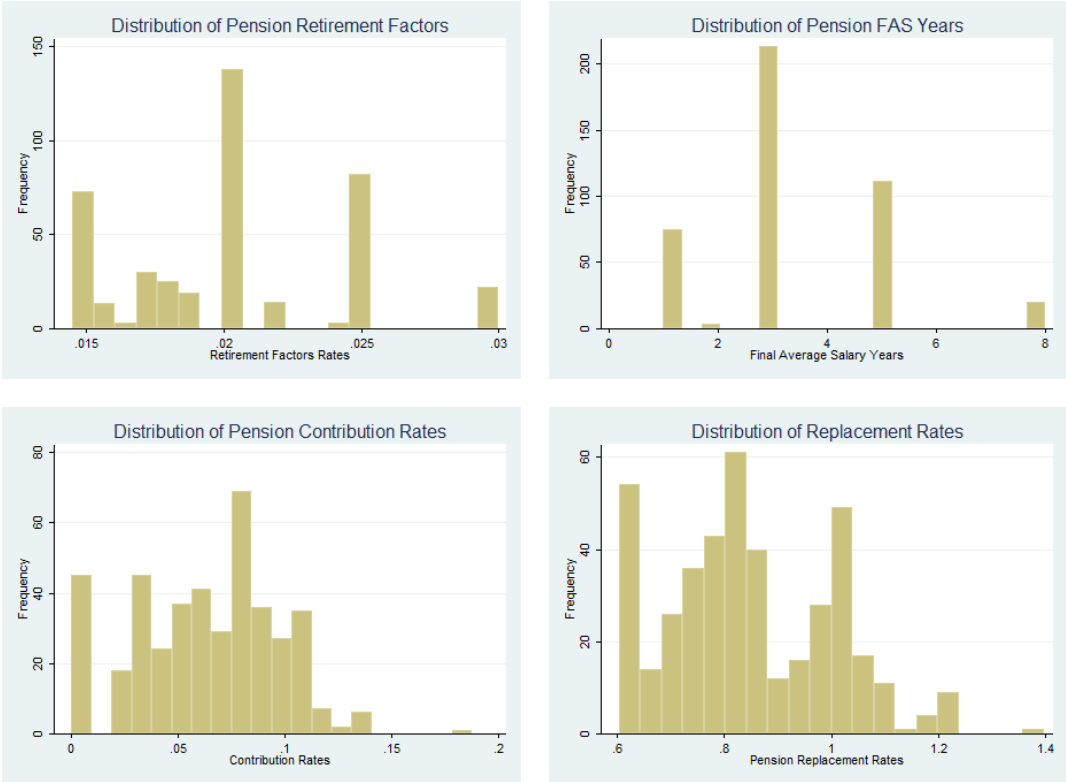
Figure 4

DISTRIBUTIONS OF PENSION PLAN PARAMETERS

4(d) shows the distribution of the estimated replacement rates. A replacement rate of one indicates the take-home pay from the pension is equal to the take-home pay in the year prior to retirement.

Many of the replacement rates in the sample tend to be larger than one, indicating that the representative public worker would receive higher take-home income in retirement than in working life. This is due in part to the fact that public sector employee pensions are sometimes exempt from state taxes, while their wages are not. Furthermore, workers must pay some of their pre-tax salary during working life in pension contributions, while they contribute nothing in retirement. These data are suggestive of high public sector pensions and backloading, in line with previous work. Beshears, Choi, Laibson, and Madrian (2011) compute replacement rates for general state employees and also finds that workers with tenure over 25 years often have replacement rates of over 100%.[24]

Finally, the degree of underfunding of pension obligations is significant and potentially important in determining the growth of pensions for public sector workers. In particular, each year an employee works, he accrues an earned benefit payable in retirement. The government is often supposed to contribute funds to the pension plan as these benefits accrue, but the degree to which this is done varies widely across states. Thus, the level of pension plans funding, defined to be total assets over total accrued liabilities, varies, with many state plans significantly underfunded. If current pension funding is too low, then there may be weaker incentives for politicians to compensate workers via pensions, as it would weaken the financial health of the pension plan and potentially necessitate higher taxes. Novy-Marx and Rauh (2009) indicate that underfunding of state-local pension plans was about $1 trillion in 2008, using discount rates of 7% and 8% on future pension liabilities. However, they argue that since future pension promises are almost universally protected strongly by contract law or state constitutions, the appropriate discount rate is the risk-

---

[24]Certainly this is not the *average* pension benefit across all public sector workers, since most workers will tend to have shorter tenures. However, given that the focus is on public sector bargaining, I focus on workers with more tenure (and therefore larger pensions) since they will likely have more power within the union.

free rate. Using Treasury borrowing rates, aggregate underfunding is $3.76 trillion for the largest 116 state pension funds. Figure 5 presents the distribution of funding ratios using the pension plans' assumed actuarial rates. All of these plans use discount rates over 7%, implying that the funding statuses of these plans is arguably weaker than indicated.

Figure 5
DISTRIBUTION OF STATED FUNDING RATIOS
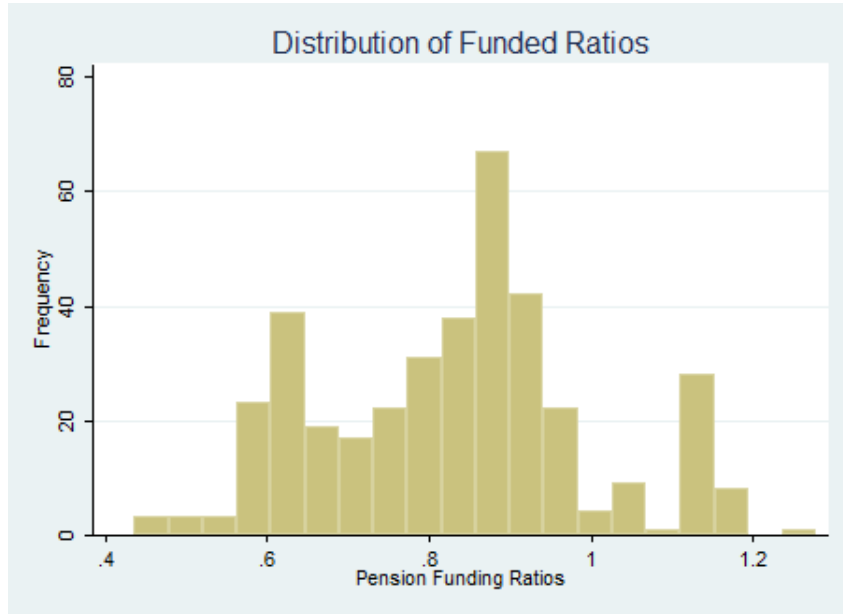
1.3.3   Data Analysis

To better understand the data, and in particular the relationship between political variables and bargaining outcomes, Table 1 shows the sample means for delay, real average wage changes, and real average pension changes, both in the aggregate and by party. The data show significant differences in all three variables by party. First, Democrats have on average significantly less delay, by about 1.5 months. Second, both real wage and real pension changes are significantly higher for Democrats. Real wage changes are approximately 0.4 percentage points larger for Democrats, while real pension changes are on average 1.3 percentage points larger. Note also that these differences are robust to state fixed effects. Overall, these data are suggestive of the potential importance of political variables in af-

fecting public sector bargaining outcomes.

Table 1

SAMPLE MEANS, BY PARTY

	All	Dem.	Rep.	Diff
Delay (in months)	6.929	6.307	7.757	-1.450*
	(0.400)	(0.486)	(0.670)	(0.828)
Avg. Real Wage	-1.833	-1.654	-2.080	0.427*
Change (in pp)	(0.125)	(0.168)	(0.185)	(0.250)
Avg. Real Pension	-1.640	-1.081	-2.410	1.328***
Change (in pp)	(0.225)	(0.264)	(0.384)	(0.476)

Figure 6

MEAN WAGES BY DELAY
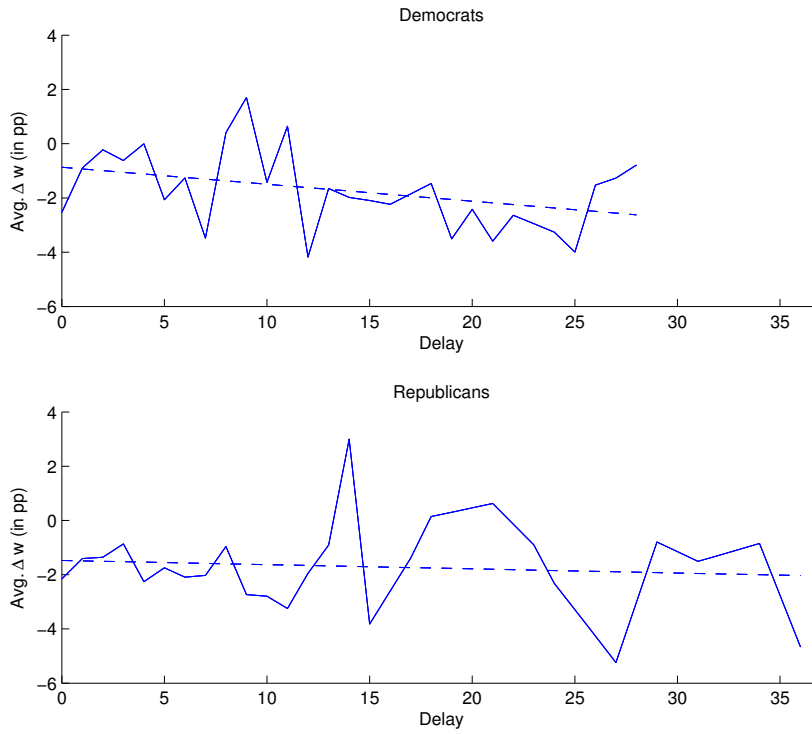
Figure 6 documents the relationship between average wage changes and delay. In partic-
ular, the graphs show, for each realized level of delay, the average of all wage changes in
contracts with this delay. The top panel plots this for Democrats while the bottom plot

if for Republicans, and the solid line is the data while the dashed is the fitted line. The data show that, for both parties, longer delay is associated with smaller wage increases. The fitted line for the Democrats indicates that a 10 month increase in delay is associated with about 0.63 percentage point lower wage growth. For Republicans, such an increase in delay is associated with about 0.16 percentage point lower wage growth. This relationship between delay and wage growth is something the model will be able to capture.

1.4   Model

To gain additional insights into the determinants of public sector union bargaining outcomes, I model the interactions between unions and governors in a stochastic bargaining framework as in Merlo and Wilson (1995) and Merlo (1997). I assume there are two players, a governor (g) and a public sector union (u). I abstract from the bargaining problem within the government (e.g. between the executive and legislature, or within the legislature) and within the union (i.e. the age conflict between young and old workers). I interchangeably refer to the union as "employees" or "workers," and to the governor as "politician." The two parties bargain over a contract that specifies the wage and pension levels for union employees. A monthly wage-pension pair is denoted $(w, p) \in \mathbb{R}^2_+$. I assume that until agreement is reached, the previous agreement, denoted $(\underline{w}, \underline{p})$, remains in place. I do not consider the ability of the union to strike, as in the sample period the presence of strikes is extremely rare, and many states explicitly forbid it.[25]

Let $\mathcal{S} = \{1, ..., S\}$ denote the (finite) set of possible states of the world with typical element s. Let σ denote a Markov process with transition probability $\pi_{s,s'} = Pr(s'|s) \geq 0$ denoting the probability of moving from state s to state s', with $\sum_{s' \in S} \pi_{s,s'} = 1$ for every $s \in \mathcal{S}$. Additionally, in each period, a proposer $\kappa \in \{g, u\}$ is selected with probability Π_κ such that

---

[25]Certainly the ability to strike may affect outcomes, even if striking is not observed in equilibrium, as long as it is a credible threat. However, the paucity of strikes does perhaps suggest that it may not be credible. For this reason, and for tractability purposes, I abstract from this dimension of bargaining.

$\Pi_g + \Pi_u = 1$. The proposal probability for each player is independent of the state $s \in \mathcal{S}$.[26]

Finally, at the beginning of each period, shock $\gamma_t \in \mathbb{R}$ is realized that affects the utility of the governor, as I will describe. This shock has CDF $H_\gamma(\gamma_t; \theta)$ with parameter θ, and is *iid* over time and independent of the state and proposer processes.

For $t = 0, ..., T$, let $\sigma^t \equiv (\sigma_0, ..., \sigma_t)$ denote the t-period state history with realization $s^t = (s_0, s_1, ..., s_t)$. That is, s^t is the $t + 1$ length history up to period t, and s_t is the realization of the state in period t. Here, $T < \infty$ is the bargaining deadline, and at $T + 1$ an election is held. Let $\tilde{\kappa}^t \equiv (\tilde{\kappa}_0, ..., \tilde{\kappa}_t)$ denote the history of proposers up to period t, with generic realization $\kappa^t \equiv (\kappa_0, ..., \kappa_t) \in \{g, u\}^{t+1}$. Finally, let $\tilde{\gamma}^t \equiv (\tilde{\gamma}_0, ... \tilde{\gamma}_t)$ denote the history of γ-shocks with generic realization $\gamma^t \equiv (\gamma_0, ..., \gamma_t) \in \mathbb{R}^{t+1}$. I assume the realized state history is observed by both the players and the econometrician, while the proposer and γ-shock histories are *only* observed the players. The wage-pension pair (w, p) is in monthly terms, and each bargaining period is one month.

The bargaining protocol is as follows. At the beginning of period t, contract $(\underline{w}, \underline{p})$ is in place and state s_{t-1} is known. Then, state s_t is realized with probability π_{s_{t-1}, s_t}, proposer $\kappa_t \in \{g, u\}$ is realized with probability Π_{κ_t}, and shock $\gamma_t \in \mathbb{R}$ is realized according to its CDF $H_\gamma(\gamma_t; \theta)$. I refer to the player who is not the proposer as the *responder*, and denote him as $-\kappa_t$. Player κ_t offers a contract (w, p), and player $-\kappa_t$ accepts or rejects. After acceptance or rejection, flow payoffs are realized. If the offer is accepted, the flow utility (and all future flow utilities) are evaluated at the new contract and the game ends. If the offer is rejected, the current period flow utility is evaluated at the old contract $(\underline{w}, \underline{p})$ and play advances one period. At time $T + 1$, a terminal payoff is realized, depending on whether agreement had ever been reached and, if so, the agreed upon contract. I assume that players are impatient with common discount factor $\delta \in (0, 1)$.

Consider a bargaining spell with initial contract $(\underline{w}, \underline{p})$. First, I describe the union's utility. Each period, it receives a flow utility depending on what the current wage and pension are.

---

[26]This assumption is not necessary for any of the main results; there may be correlation between the state and proposer.

If agreement has not been reached prior to the end of the period, the flow utility for the union is given by:

$$u_u(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) = \frac{(\underline{w}(1 - \overline{\tau}_w))^{1-\sigma}}{1 - \sigma} + \eta(X) \frac{(\underline{p}(1 - \overline{\tau}_p))^{1-\sigma}}{1 - \sigma}, \qquad (1)$$

where X is a vector of union-specific observables, $\overline{\tau}_w$ is the average tax rate during working life, $\overline{\tau}_p$ is the average tax rate on pension income in retirement.[27] The parameter $\eta(X) > 0$ governs the relative value of pensions to wages, and may capture factors like the age distribution of union workers.

If an agreement is reached, I denote the agreed upon wage-pension pair as (w^*, p^*). In a period in which agreement is reached, as well as in periods thereafter, the flow utility is:

$$u_u(w^*, p^*, X, \overline{\tau}_w, \overline{\tau}_p) = \frac{(w^*(1 - \overline{\tau}_w))^{1-\sigma}}{1 - \sigma} + \eta(X) \frac{(p^*(1 - \overline{\tau}_p))^{1-\sigma}}{1 - \sigma} \qquad (2)$$

The terminal utility if no agreement is reached by the end of time T is given by:

$$\overline{u}_u^{NA}(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) = \Upsilon u_u(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) - d_u(s_{T+1}), \qquad (3)$$

where $d_u(s_{T+1})$ is an exogenous cost to the union conditional on not reaching agreement prior to the election that depends on the economic state. The parameter Υ governs the relative value of terminal to the flow utility, and captures the notion that the contract will last several periods after the terminal period. Thus, the weight Υ captures the persistence of the contract beyond the end of bargaining.

Finally, the terminal utility if a new contract (w^*, p^*) is agreed upon before the end of

---

[27]Note that in principle, these tax rates depend on the wages and pensions of the worker due to progressive taxation. However, I use the average tax rates paid by the representative worker for $\overline{\tau}_w$ and $\overline{\tau}_p$ and assume it is a flat tax. Given the typical size in pension and wage changes, this local approximation is not quantitatively significant.

period T is given by:

$$\bar{u}_u^A(w^*, p^*, X, \bar{\tau}_w, \bar{\tau}_p) = \Upsilon u_u(w^*, p^*, X, \bar{\tau}_w, \bar{\tau}_p). \tag{4}$$

That is, the terminal utilities are affine transformations of the flow utilities from the contract.

Putting these components together, the discounted utility to the union from agreement (w^*, p^*) in state s, at time t is:

$$U_u(w^*, p^*, X, \bar{\tau}_w, \bar{\tau}_p) = \left(\frac{1 - \delta^{T-t+1}}{1 - \delta} + \delta^{T-t+1}\Upsilon\right) u_u(w^*, p^*, X, \bar{\tau}_w, \bar{\tau}_p)$$

Consider next the governor in the bargaining spell from political party $D \in \{dem, rep\}$. His utility can be broken down into four components. First, he receives a flow utility while in office throughout bargaining that depends on whether the old contract or the new contract is in place, as well as party and state s. Let $o(s, D)$ denote the flow utility when the old contract is in place, and $n(s, D)$ be the flow utility when the new contract is in place, both of which depend on the state.

Second, in the period in which agreement occurs, there is a one-time cost that depends on the state, the agreed upon contract (w^*, p^*), and the old contract. I assume a functional form of:

$$- c_1(s, D)(w^* - \underline{w}) - (c_2(s, D) - c_2^f fun)(p^* - \underline{p}), \tag{5}$$

with $c_i(s, D) \geq 0$ for all i, s, and D. Furthermore, $c_2^f \geq 0$, and fun is the funding level (i.e. assets over liabilities) of the union's pension fund. This expression captures the idea that as pension funding increases, it may become relatively less costly for the governor to increase pensions. This total cost is paid by the governor once the contract details are agreed upon and revealed, regardless of when the next election is. Third, also in the period τ in which agreement occurs, the governor receives utility γ_τ.

Finally, the governor's derives utility from the expected vote share at the election in period $T + 1$. I assume there is a unit mass of voters in the governor's state. Voter k is assumed to vote for the current politician (or his party if he is term-limited) given contract $(\hat{w}, \hat{p})$ is in place and shock s is realized on the election date, if and only if:

$$-\phi_1(\hat{w} - \underline{w}) - \phi_2(\hat{p} - \underline{p}) + \mu(s, Y) + \xi_k \geq 0,$$

with parameters $\phi_i > 0$ and $\mu(s, Y)$, and individual shock ξ_k with CDF $F(\xi_k; \lambda)$ where λ is a parameter governing the distribution. Vector Y is a set of election-specific variables that will be described further in the empirical section. The parameters ϕ_1 and ϕ_2 reflect the marginal disutility from higher wages and pensions, respectively, while $\mu(s, Y)$ reflects the competitiveness of the district. For instance, a higher value of $\mu(s, Y)$ implies that, holding fixed other terms, the incumbent party will receive a higher vote share. With a unit mass of voters, the politician's expected vote share is:

$$1 - F(\phi_1(\hat{w} - \underline{w}) + \phi_2(\hat{p} - \underline{p}) - \mu(s_{T+1}, Y); \lambda).$$

If the old contract is in place at the time of the vote, there is an exogenous cost $d_g(s_{T+1}, D)$. Thus, the expected terminal utility conditional on disagreement is:

$$1 - F(\phi_1(\hat{w} - \underline{w}) + \phi_2(\hat{p} - \underline{p}) - \mu(s_{T+1}, Y); \lambda) - d_g(s_{T+1}, D).$$

The disagreement cost here can reflect dissatisfaction voters have from gridlock, or a diminished value from being in office due to further gridlock.

In sum, the governor's utility from reaching agreement (w^*, p^*) in state s and period t is

given by:

$$U_g(w^*, p^*, s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma_t = n(s, D) + \delta \mathbb{E}_{s_{t+1}|s} n(s_{t+1}, D) + \ldots + \delta^{T-t} \mathbb{E}_{s_T|s} n(s_T, D)$$

$$- c_1(s, D)(w^* - \underline{w}) - (c_2(s, D) - c_2^f fun)(p^* - \underline{p})$$

$$+ \delta^{T-t+1} \mathbb{E}_{s_{T+1}|s}[1 - F(\phi_1(w^* - \underline{w}) + \phi_2(p^* - \underline{p})$$

$$- \mu(s_{T+1}, Y))] + \gamma_t.$$

That is, U_g is the utility of the governor excluding the unobservable component γ_t. The first component consists of the flow utilities from agreeing to a new contract. The second is the one-time cost from agreement (w^*, p^*), given funding levels. The third is the discounted utility from the expected vote share.

An outcome of this game is either: (i) a vector (τ, w^*, p^*), where $\tau \leq T$ is a stopping time (i.e. period of agreement), and (w^*, p^*) is the agreed upon wage and pension; or (ii) disagreement. The outcome implies payoffs to each party as follows. If there is agreement in τ, the discounted present value of the payoff to u is:

$$\sum_{t=0}^{\tau-1} \delta^t u_u(\underline{w}, \underline{p}, X, \bar{\tau}_w, \bar{\tau}_p) + \sum_{t=\tau}^{T} \delta^t u_u(w^*, p^*, X, \bar{\tau}_w, \bar{\tau}_p) + \delta^{T+1} \Upsilon u_u(w^*, p^*, X, \bar{\tau}_w, \bar{\tau}_p),$$

while for g it is:

$$\sum_{t=0}^{\tau-1} \delta^t o(s_t, D) + \delta^\tau \left(-c_1(s_\tau, D)(w^* - \underline{w}) - c_2(s_\tau, D)(p^* - \underline{p})\right) + \sum_{t=\tau}^{T} \delta^t n(s_t, D) +$$

$$\delta^{T+1}[1 - F(\phi_1(w^* - \underline{w}) + \phi_2(p^* - \underline{p}) - \mu(s_{T+1}, Y); \lambda)] + \gamma_\tau.$$

If instead the bargaining spell ends in disagreement, then the payoff to u is:

$$\sum_{t=0}^{T} \delta^t u_u(\underline{w}, \underline{p}, X, \bar{\tau}_w, \bar{\tau}_p) + \delta^{T+1} \left(\Upsilon u_u(\underline{w}, \underline{p}, X, \bar{\tau}_w, \bar{\tau}_p) - d_u(s_{T+1})\right)$$

22

and to g it is:

$$\sum_{t=0}^{T} \delta^t o(s_t, D) + \delta^{T+1} \left(1 - F(-\mu(s_{T+1}, Y); \lambda) - d_g(s_{T+1}, D)\right).$$

A history is a sequence of realized states, proposers, γ shocks, wage-pension offers, and acceptances or rejections in each period up to that point. A strategy for player $i \in \{g, u\}$ specifies an action (either a contract offer or an accept-reject rule) at every possible history in which he or she acts. Each possible $T+1$ length history induces an outcome and payoffs to the governor and the union. The strategies are subgame perfect if, at every possible history, each is a best response to the other player's strategy. Given that this is a finite horizon game, and that there exists an optimal strategy at each state and date there exist unique SPE payoffs. Proposition 1 formalizes this.

Proposition 1. *A SPE of this game exists, and there exists a unique set of SPE payoffs.*

Appendix 1.8.2 presents the proof of Proposition 1, as well as all other proofs. Note that although there may not be unique SPE strategies, there are unique payoffs. The reason strategies may not be unique is when the proposer is either indifferent or strictly worse off from making the best possible offer that still induces acceptance. However, in all of these equilibria the SPE payoffs are identical.

Solving the Model

Given that it is a finite game of complete and perfect information, I can solve the game for SPE payoffs and strategies using backward induction. Let $v_i(s, t, \kappa)$ denote the *ex-ante* value function for agent $i \in \{g, u\}$ in state $s \in \mathcal{S}$ and date $t \in \{1, ..., T\}$ when κ is the proposer. For notational simplicity, I abstract from dependence of the value functions on other observables. "Ex-ante" here refers to prior to the realization of the γ_t shock, but after

the state and proposer are realized. At time $T + 1$, the value functions are:

$$v_g(s, T + 1, \kappa) = 1 - F(-\mu(s, Y)) - d_g(s, D), \tag{6}$$

$$v_u(s, T + 1, \kappa) = \Upsilon \left(\frac{(\underline{w}(1 - \overline{\tau}_w))^{1-\sigma}}{1 - \sigma} + \eta(X) \frac{(\underline{p}(1 - \overline{\tau}_p))^{1-\sigma}}{1 - \sigma} \right) - d_u(s) \tag{7}$$

for each s and κ. Here, since γ_t enters additively and is mean-zero, it does not directly appear in either value function at time $T + 1$.

At time $t \leq T$, when the governor proposes in state s with shock γ_t, the optimal offer is the solution to the program:

$$\max_{(w,p) \in \mathbb{R}_+^2} U_g(w, p, s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma_t \tag{8}$$

$$\text{s.t. } U_u(w, p, X, \overline{\tau}_w, \overline{\tau}_p) = u_u(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) + \delta \mathbb{E}_{\kappa', s'|s} v_u(s', t + 1, \kappa').$$

Note first that since the governor's utility is linear in γ_t, the optimal offer does not depend on γ_t. To solve, for each w find the pension level that satisfies the constraint, denoted $p_{s,t}^g(w)$. That is, $p_{s,t}^g(\widehat{w})$ is the value of $\widehat{p}$ such that:

$$U_u(\widehat{w}, \widehat{p}, X, \overline{\tau}_w, \overline{\tau}_p) = u_u(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) + \delta \mathbb{E}_{\kappa', s'|s} v_u(s', t + 1, \kappa').$$

Given this function, the optimal w^* satisfies:

$$\max_{w \geq 0} U_g(w, p_{s,t}^g(w), s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma_t,$$

with the associated optimal pension $p_{s,t}^g(w^*).$[28]

Let $(w_{s,t}^g, p_{s,t}^g)$ denote the optimal offer by the governor at state s in time t. The governor then compares the utility from this contract with his continuation utility. There is agreement

---

[28]I show in the proof of Proposition 1 that either: (i) an optimal offer exists, or (ii) any offer the governor makes will not be accepted, and thus he either makes a rejected offer or passes. For expositional purposes, I consider the case where an optimal offer exists.

if and only if:[29]

$$\gamma_t \geq o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa') - U_g(w_{s,t}^g, p_{s,t}^g, s, t; fun, Y, D, \underline{w}, \underline{p}) \equiv \underline{\gamma}_{s,t}^g,$$

which occurs with probability:

$$1 - H_\gamma\left(\underline{\gamma}_{s,t}^g\right).$$

Note that while the cutoff rule depends on observables, I suppress this dependence for notational simplicity.

The ex-ante value functions are then:

$$v_g(s, t, g) = H_\gamma(\underline{\gamma}_{s,t}^g)\left(o(s, D) + \delta \mathbb{E}_{s'|s, \kappa'} v_g(s', t+1, \kappa')\right) +$$
$$(1 - H_\gamma(\underline{\gamma}_{s,t}^g))U_g(w_{s,t}^g, p_{s,t}^g, s, t; fun, Y, D, \underline{w}, \underline{p}) + \int_{\underline{\gamma}_{s,t}^g}^{\infty} \gamma \, dH_\gamma(\gamma; \theta)$$
$$v_u(s, t, g) = u_u(\underline{w}, \underline{p}, X, \bar{\tau}_w, \bar{\tau}_p) + \delta \mathbb{E}_{\kappa', s'|s} v_u(s', t+1, \kappa').$$

When the union proposes at time t in state s, the optimal offer now depends on the realization of γ_t. The optimal offer solves:

$$\max_{(w,p) \in \mathbb{R}_+^2} U_u(w, p, X, \bar{\tau}_w, \bar{\tau}_p) \tag{9}$$
$$\text{s.t. } U_g(w, p, s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma_t = o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa').$$

Note here that, unlike in the case in which the governor proposes, the constraint depends on γ_t. Thus, the set of pension and wages satisfying the constraint depends on the realization of the shock. That is, one can solve the constraint for optimal p as a function of w *and* γ_t, denoted $p_{s,t}^u(w, \gamma_t)$. Then, after plugging into the objective function, the optimal wage

---

[29]If this holds with equality, then it could be in equilibrium that the governor does not make an offer. As this only occurs for a measure zero of γ, this does not change any of the results, and thus I assume if the proposer is indifferent, he makes the offer.

offer, given γ_t, satisfies:

$$\max_{w \geq 0} \ U_u(w, p^u_{s,t}(w, \gamma_t), X, \overline{\tau}_w, \overline{\tau}_p)$$

Although in this case the function $p^u_{s,t}(w, \gamma_t)$ cannot be solved for in closed form, it can be solved for numerically. Furthermore, the implicit function theorem gives the derivative of this function with respect to w, which is needed when taking the first-order condition.

The following lemma states that, when proposing, the union follows a cutoff rule. That is, the union proposes if and only if γ_t is sufficiently large. Intuitively, if γ_t increases, then the governor is more inclined to accept this period, and therefore the union can extract more of the surplus, if there is any.

Leamma 1. *The union follows a cutoff rule such that, for all $\gamma_t \geq \underline{\gamma}^u_{s,t}$, an offer is made and accepted, and for all $\gamma_t < \underline{\gamma}^u_{s,t}$ the union passes.*

The following lemma argues that in all states and dates, the cutoff rule is independent of the proposer.

Leamma 2. *In any SPE, for all s and t, $\underline{\gamma}^g_{s,t} = \underline{\gamma}^u_{s,t}$.*

Given this, let $\underline{\gamma}_{s,t}$ denote the cutoff shock at time t and state s. Furthermore, I denote the optimal offer by the union in state s, time t, and with shock γ by $(w^u_{s,t}(\gamma), p^u_{s,t}(\gamma))$. The ex-ante value functions in period t and state s when the union proposes are then given by:

$$v_g(s, t, u) = o(s, D) + \delta \mathbb{E}_{s'|s,\kappa'} v_g(s', t+1, \kappa')$$

$$v_u(s, t, u) = H_\gamma(\underline{\gamma}_{s,t}) \left(u_u(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) + \delta \mathbb{E}_{\kappa', s'|s} v_u(s', t+1, \kappa') \right) + \int_{\underline{\gamma}_{t,s}}^{\infty} U_u(w^u_{s,t}(\gamma), p^u_{s,t}(\gamma), X, \overline{\tau}_w, \overline{\tau}_p) \, dH_\gamma(\gamma; \theta).$$

Furthermore, conditional on reaching period t without agreement, the probability of agreement at t in state s is given by $1 - H_\gamma\left(\underline{\gamma}_{s,t}\right)$. Given a history of states s^T, the probability

that delay is zero (again, implicitly given observables) is $1 - H_\gamma\left(\underline{\gamma}_{s_0,0}\right)$. The probability of a delay of length $1 \leq N \leq T$ is:

$$\left(\prod_{n=0}^{N-1} H_\gamma\left(\underline{\gamma}_{s_n,n}\right)\right) \times \left(1 - H_\gamma\left(\underline{\gamma}_{s_N,N}\right)\right),$$

while the probability of disagreement, which is a delay of $T + 1$, is:

$$\prod_{n=0}^{T} H_\gamma\left(\underline{\gamma}_{s_n,n}\right).$$

Similarly, the model will imply a distribution of wages and pensions.

The set of ex-ante value functions can be solved by backward induction, similar to the theoretical model. As the integrals are not in closed form, I use Monte Carlo integration techniques to compute them.

Empirical Specification

I now describe the empirical specification of the model. I assume that the structure of the γ shocks is as follows. In bargaining spell b in geographic state j at time t, the γ shock is $\gamma_{t,b,j} = \gamma^1_{t,b,j} + \gamma^2_{t,j}$, where $\gamma^1_{t,b,j} \sim \mathcal{N}(0, \theta_1)$ and $\gamma^2_{t,j} \sim \mathcal{N}(0, \theta_2)$, both iid over time and independent of each other. That is, the shock is composed of an idiosyncratic component and a component that is common to all unions in the same state that are bargaining at the same time. This captures the common factors of bargaining and helps match the observed correlations in outcomes among unions in the same state bargaining at the same time. For the remainder of this section, I denote the sum of the idiosyncratic and geographic state-specific shocks as γ_t, suppressing dependence on bargaining unit and geographic state. The CDF γ_t is then $N_\gamma(\cdot; \theta_1 + \theta_2)$, with $\theta_1 + \theta_2$ as the variance of the shock. Note that this does not change any of the results from the previous section, and the shock is still observed by both players, but not the econometrician.

Next, I assume in empirical implementation that $X = \{geo.\ state, PS\}$. That is, the

weight on pension utility for the union, $\eta(X)$, depends on the geographic state and PS, an indicator for public safety unions. This specification captures the large differences in pension size between public safety and non-public safety unions, likely due in part to the fact that public safety workers tend to retire earlier due to more strenuous work, and thus may value pensions more than wages relative to non-public safety workers. Furthermore, the dependence on geographic state helps match differences across different geographic states.

I assume that $Y = \{OS, \widetilde{CPVI}\}$, where OS is a dummy variable for if the next election is an open seat election. This captures the effect of incumbency on expected vote share. The variable $\widetilde{CPVI}$ is the Cook Partisan Voting Index in favor of the governor's party at the beginning of bargaining.[30] As described in the data section, the standard CPVI measures the degree to which the district leans Democrat or Republican according to previous presidential election vote shares. I adjust this measure so that $\widetilde{CPVI}$ is positive if the district leans in the same direction as the governor's party, and is negative otherwise. This captures overall district partisanship and affects the expected vote share.

I make the following functional form assumptions and normalizations for empirical implementation. First, I assume for simplicity that $o(s, D) = 0$ for all s and D. I assume that all parameters depending on the state are linear in s. Explicitly, for each party $D \in \{dem, rep\}$, I assume there are two parameters n_1^D and n_S^D. Then, I let the benefit to a governor from being under a new contract be:

$$n(s, D) = n_1^D + \frac{s-1}{S-1}(n_S^D - n_1^D), \text{ for all } s \in \{1, ..., S\}.$$

<hr>

[30]The CPVI can in principle change if a bargaining spell overlaps with presidential election. In 11 states, gubernatorial elections are typically held in the same year as presidential elections. In three states, they are 12 months prior; in 36 states they are 24 months prior; and in two states, they are 36 months prior. Note this sums to 52, since New Hampshire and Vermont hold elections every two years. Thus, delay must be quite long for the CPVI to change. Furthermore, the CPVI does not typically change significantly over different presidential elections, especially within sample. Since allowing it to vary stochastically would increase the state space and increase the time to solve the model with only minor benefits, I abstract from varying CPVI within a bargaining spell.

The terminal disagreement cost for the governor $d_g(s, D)$ takes the form:

$$d_g(s, D) = d_{g,1}^D + \frac{s-1}{S-1}(d_{g,S}^D - d_{g,1}^D), \text{ for all } s \in \{1, ..., S\}$$

with parameters $\{d_{g,1}^D, d_{g,S}^D\}_{D \in \{dem,rep\}}$. The terminal disagreement cost takes the form:

$$d_u(s) = d_{u,1} + \frac{s-1}{S-1}(d_{u,S} - d_{u,1}), \text{ for all } s \in \{1, ..., S\}$$

with parameters $d_{u,1}$ and $d_{u,S}$.

The one-time costs for the governor from agreement depend on the state and the party. Both $c_1(s, D)$ and $c_2(s, D)$, and take the forms:

$$c_i(s, D) = c_{i,1}^D + \frac{s-1}{S-1}(c_{i,S}^D - c_{i,1}^D) \text{ for } D \in \{dem, rep\}, \ s \in \{1, ..., S\}, \text{ and } i \in \{1, 2\},$$

with parameters $\{c_{1,1}^D, c_{1,S}^D, c_{2,1}^D, c_{2,S}^D\}_{D \in \{dem,rep\}}$.

I assume that $\mu(s, Y)$ takes the form:

$$\mu(s, Y) = \mu_1 + \frac{s-1}{S-1}(\mu_S - \mu_1) + \chi_1\widetilde{CPVI} + \chi_2 OS, \text{ for all } s \in \{1, ..., S\},$$

with parameters μ_1, μ_S, χ_1, and χ_2. I also assume ξ_k is exponentially distributed with parameter λ.

In specifying the union's relative benefit from pension, I assume that $\eta(X) = \eta^j + \eta^{ps,j} PS$ for geographic state j. The parameters η^j are the geographic state-specific relative value of pensions to wages, and $\eta^{ps,j}$ are the (also geographic state-specific) additional value for public safety unions. This reflects the fact that public safety employees typically retire earlier and value pensions more. It also is to match the empirical regularity that public safety employees have typically larger pensions. The heterogeneity in the value of pensions also captures differences across states, including the age-structure of public sector employees.

In stochastic bargaining models the proposal probability partially determines the bargaining power of each side. Intuitively, public sector workforces with stronger unions may have more bargaining power. Thus, I assume the probability the union proposes depends on the percentage of public sector workers in a union, denoted MEM, and is given by:

$$\Pi_u(MEM) = \overline{\Pi}_u + \beta_1 MEM.$$

Finally, I specify the state space as follows. For each geographic state, I compute the mean unemployment rate since 1976. Then, for each monthly observation of unemployment, I take the deviation from the state average and denote it $UE_{j,t}$ for geographic state j in month t. I then define discrete bins with cutoffs $\{\overline{UE}_1, ..., \overline{UE}_{S-1}\}$, independent of the geographic state, such that state $s = 1$ corresponds to $UE_{j,t} \leq \overline{UE}_1$, states $s \in \{2, ..., S-1\}$ correspond to $UE_{j,t} \in (\overline{UE}_{s-1}, \overline{UE}_s]$, and state $s = S$ corresponds to $UE_{j,t} > \overline{UE}_{S-1}$. This then defines, for each geographic state and month, a state in the Markov process. In empirical implemetation, I assume $S = 8$, with $\{\overline{UE}_1, \overline{UE}_2, ..., \overline{UE}_7\} = \{-2.5, -1.5, ..., 3.5\}$.

This fully specifies the empirical model. Letting $\mathcal{G}$ denote the set of geographic states, the model parameters are:

$$\Theta \equiv \{\{\{\pi_{r,s}\}\}_{r,s=1}^{S}, \{\{d_{g,s}^D, n_s^D, c_{1,s}^D, c_{2,s}^D\}_{s \in \{1,S\}}\}_{D \in \{dem,rep\}}, c_2^f, \phi_1, \phi_2, \lambda, \chi_1, \chi_2,$$
$$\{\mu_s, d_{u,s}\}_{s \in \{1,S\}}, \{\eta^j, \eta^{ps,j}\}_{j \in \mathcal{G}}, \Upsilon, \delta, \sigma, \theta_1, \theta_2, \overline{\Pi}_u, \beta_1\}.$$

Given that $S = 8$ and $\#\mathcal{G} = 17$, in total there are 131 parameters: 64 for the transition probabilities, 34 for the weights on pensions in the union's utility, and 33 other parameters. I normalize the initial wage $\underline{w}$ to 1, and all other variables are relative to the initial salary.

1.5 Calibration and Fit

In this section I discuss the calibration of the model and document the model fit. First, I estimate $\{\{\pi_{r,s}\}\}_{r,s=1}^{S}$ outside of the model from the observed transition probabilities in the data. Table 2 shows the estimated transition matrix.

Table 2
PROBABILITY OF TRANSITION, STATE i TO j

i,j	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
s_1	0.9239	0.0761	0	0	0	0	0	0
s_2	0.0221	0.9313	0.0456	0.0008	0	0	0	0
s_3	0.0002	0.0029	0.9305	0.0395	0.0003	0.0002	0	0
s_4	0	0.0002	0.0461	0.9153	0.0380	0.0004	0	0
s_5	0	0	0	0.0659	0.8935	0.0400	0.0006	0
s_6	0	0	0	0	0.0716	0.8833	0.0441	0.0009
s_7	0	0	0	0	0	0.0927	0.8622	0.0451
s_8	0	0	0	0	0	0	0.0612	0.9368

For δ, I use the monthly discount factor of 0.99 and I set $\sigma = 3$, both consistent with many estimates from the literature (see Scholz, Seshadri, and Khitatrakun (2006), for example, which uses $\sigma = 3$ and $\beta = 0.96$ for a model with one-year periods). Given these, I assume that $\Upsilon = 100$ since it is approximately $\frac{1}{1-\delta}$, which would be the weight for an infinitely lived union.[31]

I set the voter preference parameters $\Theta_{vot} = \{\{\mu_s\}_{s\in\{1,S\}}, \phi_1, \phi_2, \lambda, \chi_1, \chi_2\}$ to match, given the observed contract wage and pension (w^*, p^*), initial wage and pension $(\underline{w}, \underline{p})$, terminal state s_{T+1}, and vote shares. More precisely, for bargaining spell b, let VD_b the vote difference between the incumbent's party and the opposition party in the following election as observed in the data. The implied vote difference from the model, given parameters, wages w_b^* and $\underline{w}_b$, pensions p_b^* and $\underline{p}_b$, observables $\widetilde{CPVI}_b$ and OS_b, and terminal state s_{T_b+1},

---

[31]Note that the results are not particularly sensitive to these parameters.

is:

$$\overline{VD}_b(\Theta_{vot}; w_b^*, \underline{w}_b, p_b^*, \underline{p}_b, s_{T_b+1}, \widetilde{CPVI}_b, OS_b) =$$

$$1 - 2F(\phi_1(w_b^* - \underline{w}_b) + \phi_2(p_b^* - \underline{p}_b) - \mu(s_{T_b+1}, \widetilde{CPVI}_b, OS_b); \lambda).$$

I then pick the voting preference parameters Θ_{vot} to minimize:

$$\sum_b (VD_b - \overline{VD}_b(\Theta_{vot}; w_b^*, \underline{w}_b, p_b^*, \underline{p}_b, s_{T_b+1}, \widetilde{CPVI}_b, OS_b))^2.$$

Note if the bargaining spell b ends in disagreement, I let $w_b^* = \underline{w}$, and $p_b^* = \underline{p}$.

I set η^j and $\eta^{ps,j}$ as follows. Consider a union in state j facing tax rates $\overline{\tau}_w$ and $\overline{\tau}_p$ whose problem is to maximize their "steady state" utility:

$$\frac{(w(1 - \overline{\tau}_w))^{1-\sigma}}{1 - \sigma} + \widehat{\eta}_b \frac{(p(1 - \overline{\tau}_p))^{1-\sigma}}{1 - \sigma}$$

subject to allocating a fixed amount of income between wages and pensions,[32] given a weight $\widehat{\eta}_b$ on pensions. That is, the solution to such a maximization problem would give the optimal wage-pension ratio. The solution to this problem would imply:

$$\widehat{\eta}_b = \frac{p^\sigma (1 - \overline{\tau}_w^j)^{1-\sigma}}{w^\sigma (1 - \overline{\tau}_p^j)^{1-\sigma}}.$$

Therefore, I set η^j to be the average value of $\widehat{\eta}_b$ among all non-public safety unions in state j, evaluated at $(\underline{w}, \underline{p})$. I set $\eta^{ps,j}$ to be the be such that $\eta^j + \eta^{ps,j}$ is the average value of $\widehat{\eta}_b$ among all public safety unions in state j.

I normalize $\overline{\overline{\Pi}}_u = 0.45$ and $\beta_1 = 0.1$ since intuitively one may expect higher union membership rates to have more bargaining power. It also implies that a state in which membership is $MEM_b = 0.5$ will have $\overline{\overline{\Pi}}_u + \beta_1 MEM_b = 0.5$. I set $\theta_1 = \theta_2$, and then calibrate $\theta \equiv \theta_1 + \theta_2$ to match the observed spread in wages and pensions given the other parameters.

---

[32]The amount to be allocated does not matter given the preferences.

The remaining parameters are flow benefits $n(s, D)$, one-time agreement costs $c_1(s, D)$, $c_2(s, D)$, and c_2^f, and disagreement costs $d_u(s)$ and $d_g(s, D)$. For intuition on how these distinctly affect bargaining, consider again the governor's utility from agreement:

$$
\begin{aligned}
n(s, D) &+ \delta \mathbb{E}_{s_{t+1}|s} n(s_{t+1}, D) + ... + \delta^{T-t} \mathbb{E}_{s_T|s} n(s_T, D) \\
&- c_1(s, D)(w^* - \underline{w}) - (c_2(s, D) - c_2^f fun)(p^* - \underline{p}) \\
&+ \delta^{T-t+1} \mathbb{E}_{s_{T+1}|s} [1 - F(\phi_1(w^* - \underline{w}) + \phi_2(p^* - \underline{p}) - \mu(s_{T+1}, Y))] + \gamma_t.
\end{aligned}
$$

The first line corresponds to the flow benefits. Note that when bargaining is in a period far from T, then this involves a sum of a potentially large number of terms. Thus, if flow benefits are relatively high, in early periods there will be more pressure for the governor to reach agreement all else equal. This will primarily affect both the hazard rate of bargaining spells and the decline in wages as delay increases. The second line corresponds to the one-time costs. These primarily affect the relative levels of wage and pension growth. The third line consists of the discounted benefit from the vote share and the γ_t shock, which have already been discussed. The final component, which does not appear in the utility from agreement, are the disagreement costs. These tend to affect the probability of disagreement especially among bargaining spells that last close to T, since early in bargaining it is discounted relatively heavily. They also affect the levels of compensation throughout bargaining. Thus, these remaining parameters all tend to have stronger effects on different dimensions of bargaining outcomes.

To calibrate $d_u(s)$ and $d_g(s, D)$, I calibrate in three steps. First, to set the difference between $d_{u,1}$ and $d_{u,S}$, I consider the following. Since these parameters in part reflect the future value of continuing bargaining after the election, note that if the union in bargaining spell b disagrees, his utility is:

$$
\Upsilon \left(\frac{(\underline{w}(1 - \bar{\tau}_w))^{1-\sigma}}{1 - \sigma} + \eta(X) \frac{(\underline{p}(1 - \bar{\tau}_p))^{1-\sigma}}{1 - \sigma} \right) - d_u(s).
$$

To calibrate $d_{u,S} - d_{u,1}$, I first for $d_{u,S}$ take the difference between the average utility from observed agreements with spells beginning in state S (given preference parameters) and the average utility from initial contracts. Then, for $d_{u,S}$ I calculate the same for bargaining spells beginning in state 1. I finally scale these by Υ and take the difference.

Second, for $d_{g,S}^D - d_{g,1}^D$, I consider spells that last until the final bargaining period. Given the values of other parameters, I compute the values of $d_{g,S}^D$ and $d_{g,1}^D$ that imply the probability of disagreement observed among these spells in the data by party. The difference between these two values gives the calibrated level of $d_{g,S}^D - d_{g,1}^D$. Finally, to calibrate the *levels* of the disagreement costs, I jointly adjust them so that, given other parameters, the model roughly matches both the average levels of compensation and overall probability of disagreement in the full sample.

This leaves $\{\{\{n_s^D, c_{1,s}^D, c_{2,s}^D\}_{s \in \{1,S\}}\}_{D \in \{dem,rep\}}, c_2^f\}$. I calibrate the flow benefits to roughly match average delay, the hazard rate of bargaining spells, and the decline in wages as delay increases, each by party. I calibrate the one-time costs $c_{1,s}^D$, $c_{2,s}^D$, and c_2^f to roughly match the relative values of wages and pensions by party, and the responsiveness of pension growth to changes in funding. Table 5 in appendix 1.8.1 presents the calibrated parameter values.

1.5.1   Model Fit

Given these calibrated parameters, I solve the model and then generate model simulations for the various bargaining spells in the data. Fixing the parameters and given the set of observables for each bargaining spell, I solve the model using backward induction as discussed above. Then, I draw a realized sequence of $\{\{\gamma_{t,b,j}^1\}_b, \gamma_{t,j}^2\}\}_{t,j}$ and, using the true sequence of unemployment rates, I use the decision rules to find a realized delay time, wage, and pension agreement, or disagreement if none is reached. This gives a distribution of agreements and delays that can be compared to the data to see how well the model is performing.

For these simulations, in the initial period, I draw *two* values for $\{\{\gamma_{0,b,j}^{1}\}_{b}, \gamma_{0,j}^{2}\}\}_{j}$ each. I assume that the maximum of each shock is the realized draw for the shock, and after observing all shocks, both sides make their offers and decisions. This is due to the fact that, in reality, bargaining may occur multiple periods before the begin date of the contract. This is not explicitly modeled because I do not have information on when bargaining between the governor and union actually begins for all states. It also helps to match the large number of bargaining spells with no delay. Note that this adjustment does not affect any of the cutoff rules in any state or date, since it occurs in the initial period and the solution method involves backward induction. It simply increases the probability of there being agreement in the initial period.

Table 3 presents the sample means for delay, net wage changes, and net pension changes for the data and simulations, both in the aggregate and by party. Furthermore, Figures 7 and 8 show the distributions of net wage and pension changes conditional on agreement, respectively. The top graphs in each panel show the distribution of simulated output, while the bottom graphs present the data. The model generates sample means for the outcome variables close to those in the data. The mean net wage change in both the data and the simulations is approximately -1.83 percentage points. For net pension changes, the mean in the data is -1.64 percentage points, while in the simulations it is -1.73 percentage points. Additionally, the model only slightly overstates average delay. In the data, mean delay is approximately 6.93 months, whereas in the simulations it is 7.07 months. Thus, the model does relatively well in broadly matching the average compensation changes and delay in the data.

Looking more closely at the unconditional distributions of wages and pensions, I note several features. First, while the model broadly captures the shapes and ranges of the distributions, the distribution of simulated wages is more concentrated than that of the data. However, the standard deviations are similar: in the data it is about 2.44, whereas in the simulations it is 2.20. The distribution of pension changes are more noticeably different, namely in that, in the data, there are some extreme values. However, outside of these values, the

35

Table 3

Table 3
SAMPLE MEANS, DATA AND SIMULATION

	Data				Simulated			
	All	Dem.	Rep.	Diff	All	Dem.	Rep.	Diff
Delay	6.929	6.307	7.757	-1.450*	7.066	6.494	7.829	-1.335
(in mos.)	(0.400)	(0.486)	(0.670)	(0.828)	(0.435)	(0.532)	(0.724)	(0.898)
Avg. Δw	-1.833	-1.654	-2.080	0.427*	-1.825	-1.604	-2.135	0.531**
(in pp)	(0.125)	(0.168)	(0.185)	(0.250)	(0.109)	(0.136)	(0.175)	(0.222)
Avg. Δp	-1.640	-1.081	-2.410	1.328***	-1.726	-1.313	-2.306	0.994***
(in pp)	(0.225)	(0.264)	(0.384)	(0.476)	(0.078)	(0.105)	(0.101)	(0.145)

distribution of data is slightly less dispersed around the mean than in the simulations. The standard deviations of net pension changes are also noticeably different, at 4.44 for the data and 1.60 for the simulations, driven largely in part by the outliers.

Figure 7
DISTRIBUTIONS OF NET WAGE CHANGES, SIMULATED VS. DATA
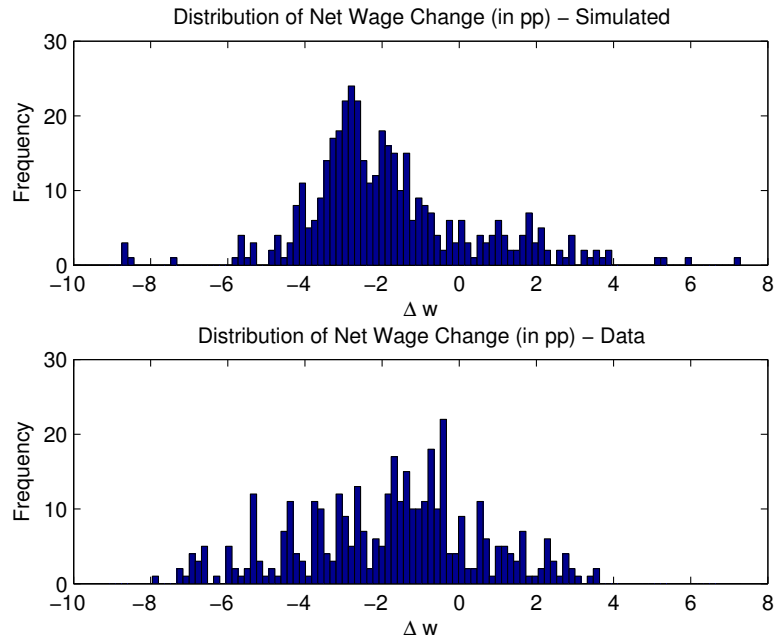
The model also predicts delay times. Figure 9 shows the distributions of delay in the simulations versus the data. The model does a relatively good job at matching the broad

Figure 8
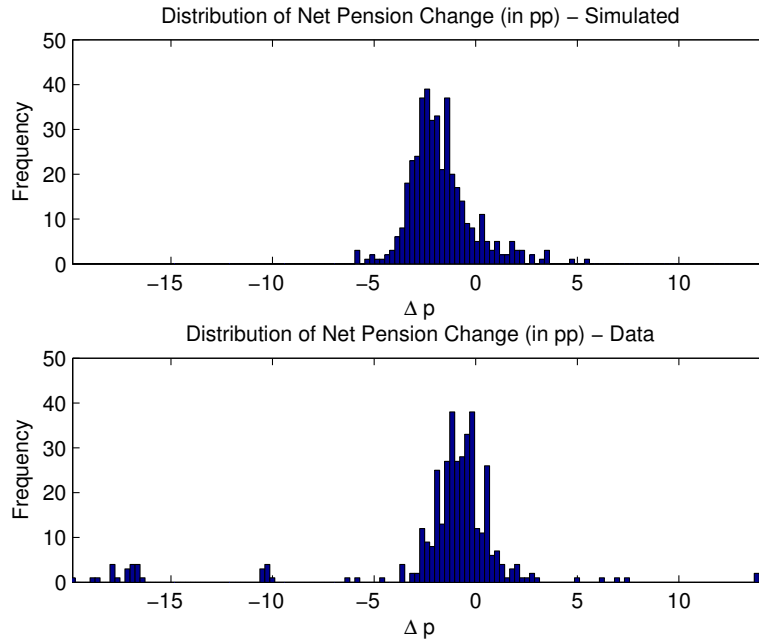
patterns in the data. Note the large spike at a delay time of zero in both the simulations and the data, corresponding to the large number of bargaining spells that reach agreement before the contract begins. This mean of the distribution of delay in the data is 6.93 months, while the simulations it is 7.07. The standard deviations of delay are 8.23 in the data and 8.93 in the simulations.

The distributions can also be broken down by party of the governor. Figure 10 shows the simulated wages and pension changes by party. The top panel shows net real wage changes and the bottom panel shows net real pension changes. The top row of each panel shows simulated output and the bottom row shows the data. Within each panel, the left column shows the results for Republican governors, and the right does so for Democrats. In the simulations, the mean pension change for Republicans is -2.31 percentage points, whereas for Democrats it is -1.31. In the data, it is -2.41 percentage points for Republicans and -1.08 for Democrats. The model captures the fact that pension changes are smaller under Republicans. For wages, under Republicans the average wage growth conditional

Figure 9
DISTRIBUTIONS OF DELAY, SIMULATED VS. DATA

38
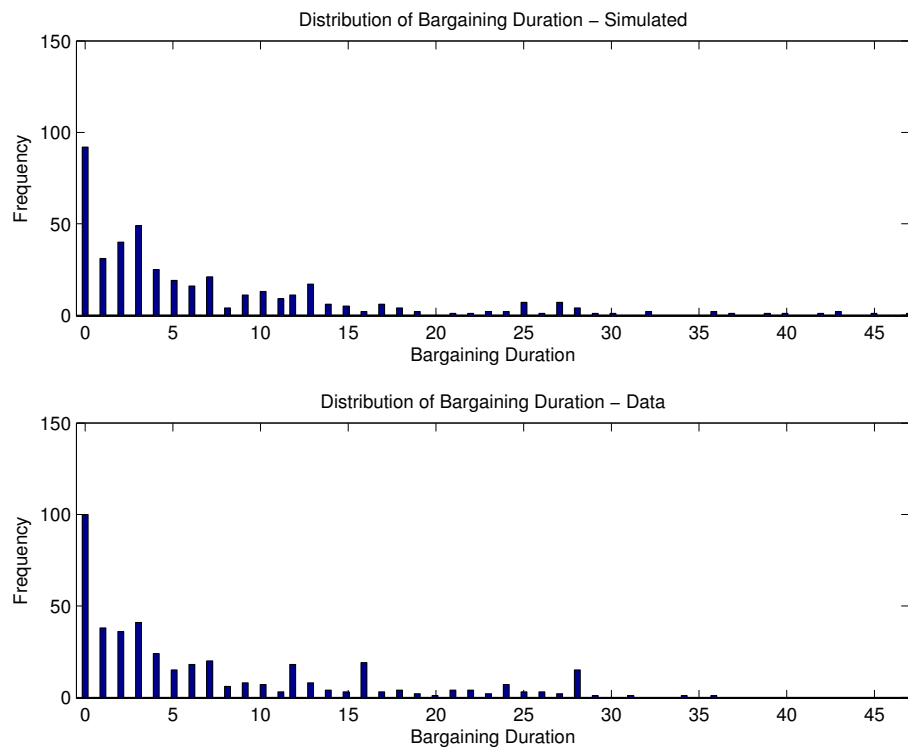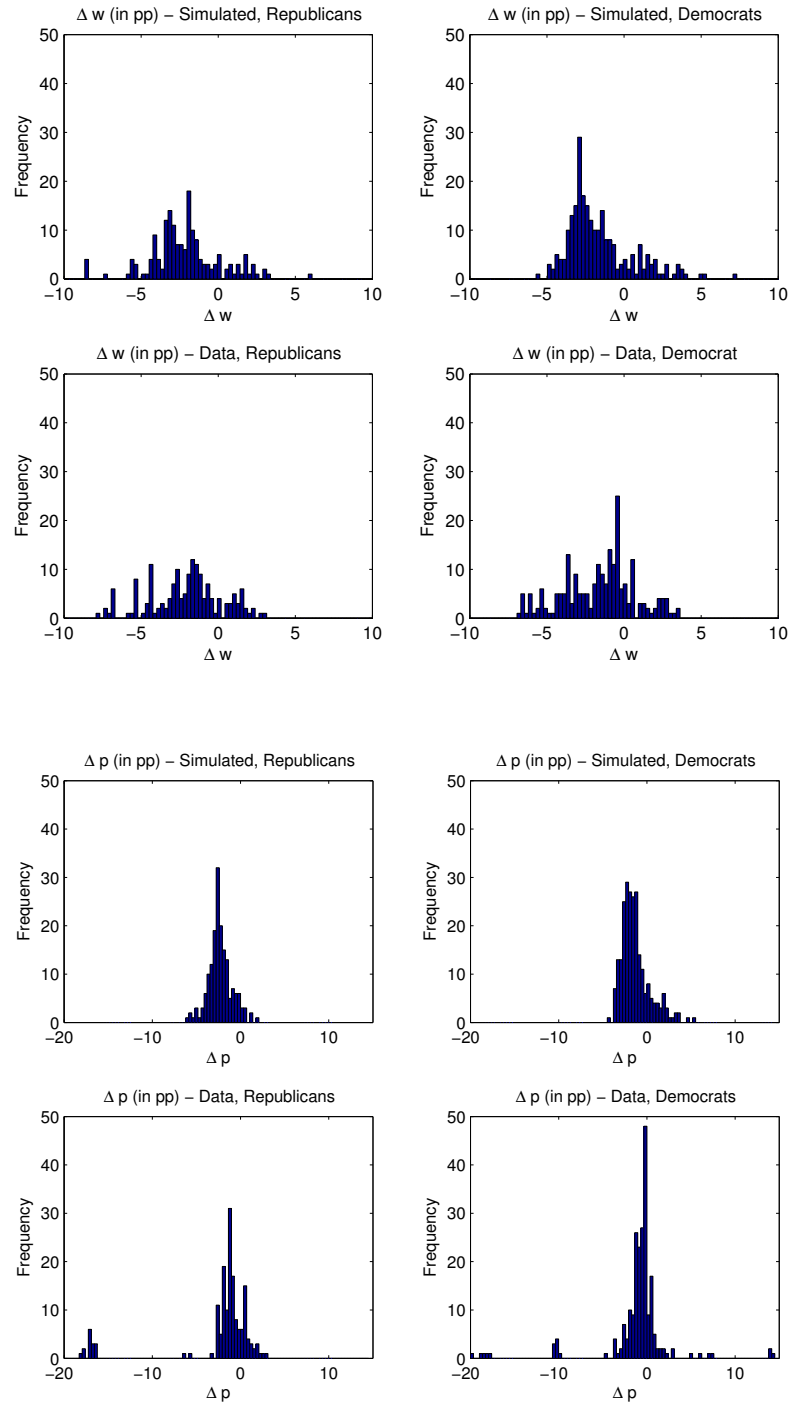
Figure 10

Distributions of Simulated Model Output, By Party

on agreement in the simulations is -2.14 percentage points, while in the data it is -2.08. For Democrats, the average simulated wage is -1.60, while the average in the data is -1.65. Again, the model accurately captures the differences in wage increases by party, in particular that wage growth is smaller under Republicans. As noted in Table 3, the differences across parties for wage and pension changes are significant in both the data and the simulations.

Figure 11 shows the delay distributions by party. The mean delay in the data for Democrats is 6.31 months, whereas for Republicans it is 7.76. In the simulations, mean delay for Democrats is 6.49 months, while for Republicans it is 7.83. Furthermore, the distributions of delay by party in the simulations are relatively similar to those in the data. For Democrats, the standard deviation of the distribution in the data is 7.54, while in the data it is 8.25. For Republicans, it is 9.01 in the data and 9.73 in the simulations. Thus, the model captures the relative differences in delay by party observed in the data.

Finally, I show that the model captures the negative relationship between delay and average wage changes observed in the data. Similar to Figure 6, the plots in Figure 12 show the average wage change among all contracts with a given level of delay. The top panel is for Democrats, while the bottom is for Republicans. Within each panel, the blue line plots the average wage change for each level of delay in the data, while the red line plots the same for the simulations. The blue and red dashed lines are the fitted lines for the data and simulations, respectively. Note that the model captures the observation in the data that average wage changes tend to be smaller in bargaining spells featuring longer delay. For Democrats, the slope of the fitted line for the data is approximately -0.063, while in the simulations it is -0.064. This implies that in the sample, bargaining spells with 10 months longer delay have approximately smaller wage increases by approximately 0.6 percentage points. The simulations match this feature well. For Republicans, the model implies slightly larger sensitivity of average wage changes to delay. The slope of the fitted line in the data is -0.016, while in the simulations it is -0.041. Still, overall the model does capture this negative relationship observed in the data.

40

Figure 11

DISTRIBUTIONS OF DELAY, SIMULATED VS. DATA, BY PARTY
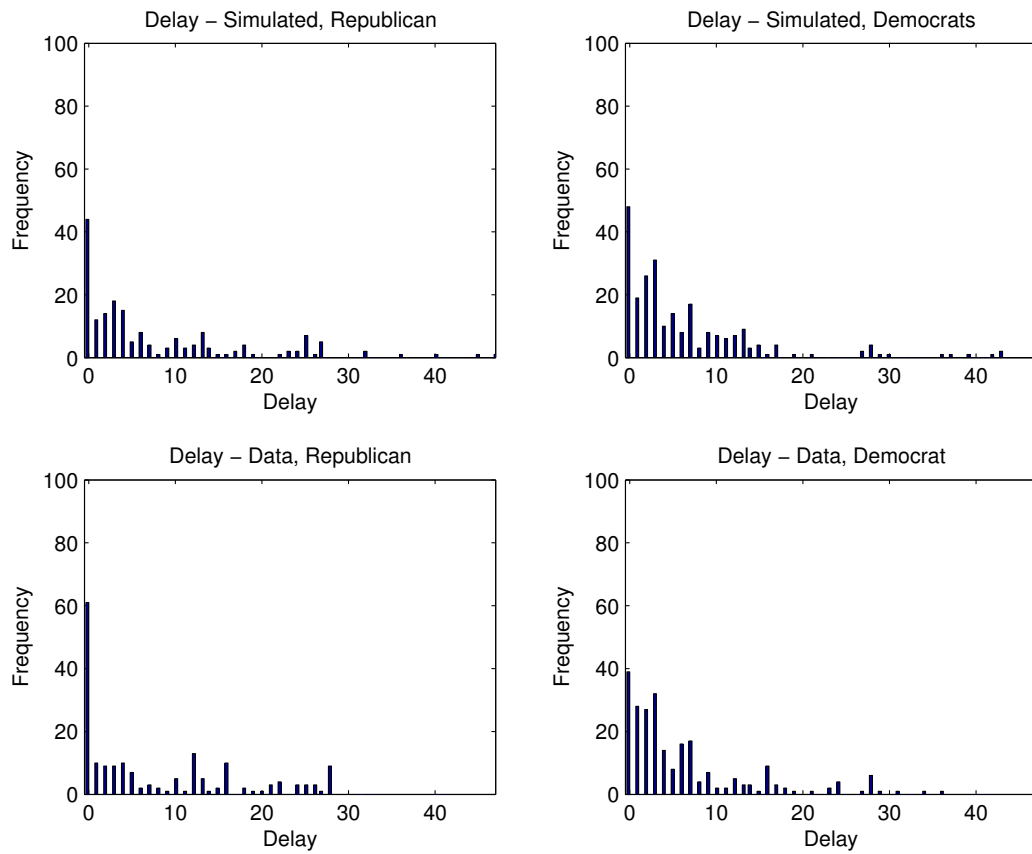
Figure 12

MEAN WAGES BY DELAY
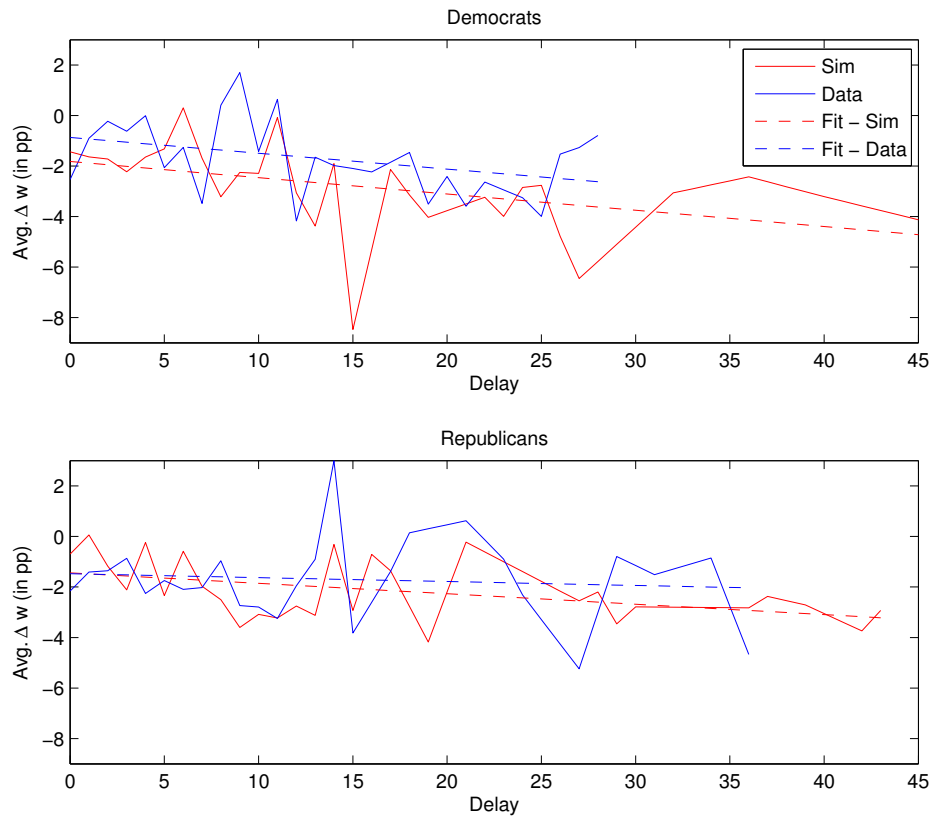
1.6 Counterfactuals

To further investigate the implications of the model, I run several counterfactuals to investigate the importance of different political and economic factors on bargaining outcomes. First, to understand the impact of the governor's party on the various bargaining outcomes, I consider a bargaining spell with median median levels of union membership, taxes, and $\widetilde{CPVI}_i$. I also assume the spell involves a non-public safety union, and take the median values of η^j and $\underline{p}$ in the sample. Given the values and the calibrated parameters, I solve the game varying party and T. Then, with the model solutions, I simulate realizations of the economics states in $\mathcal{S}$, as well as the shocks $\{\{\gamma_{t,b,j}^1, \gamma_{t,j}^2\}\}$.[33] I also vary the initial state for the shock across these simulations to understand the effect of initial unemployment on bargaining outcomes.

The results for the average of $N = 1000$ simulations for delay, wage changes, and pension changes are plotted in Figures 13, 14, and 15 respectively. The top left panel plots outcomes for Democrats, while the top right plots them for Republicans. The bottom panel plots Democrats minus Republicans, so a negative number indicates Democrats have shorter delay than Republicans. For each panel, length of bargaining spell T is on the x-axis, and the y-axis is delay in months. Each different line in the plots corresponds to the average for a different initial state s_0.

As in the aggregate statistics, Democrats generally have a shorter delay times than Republicans, with interesting non-linear effects. For instance, for bargaining spells beginning 40 months prior to the next election, average delay is about 4 months longer when the initial state is bad, while it is only about 0.5 months longer in a good state. For a bargaining length of $T = 16$, the difference is generally around 1 month for good and bad states, and only about 0.2 to 0.4 months for average states.

Finally, note that the response of delay to changes in T depends strongly on the initial state

---

[33]Note that here the distinction between the two shocks is not relevant, as I only consider one election in a given state.

s_0. The intuition for this is the following. In good states, the cost of delay is relatively high since the flow benefits conditional on agreement $n(s, D)$ are high. Therefore, if the governor chooses to disagree early in bargaining, there is a positive probability he will have to go several periods with low flow benefits $o(s, D) = 0$. Furthermore, the larger T is, the higher is the potential number of periods with low flow benefits. Thus, regardless of T, there is strong pressure early in bargaining to reach agreement – in fact, it can be stronger for larger T, which explains why, when $s_0 = 2$, average delay in fact declines slightly as T grows. However, in bad states $n(s, D)$ is low, potentially even negative. Therefore, when T is large, there is little pressure to reach agreement since disagreement costs are far in the future. The union and governor can mutually benefit from waiting for a better state or γ shock to come, so long as they do not get too close to the terminal date. Instead, for small T both sides are already close to the disagreement costs, so they tend to reach agreement quickly. Thus, the dynamics are in part driven by the relative importance of flow and disagreement costs, which depend on the initial state and T.

Figure 13
AVERAGE DELAY, SIMULATIONS

44
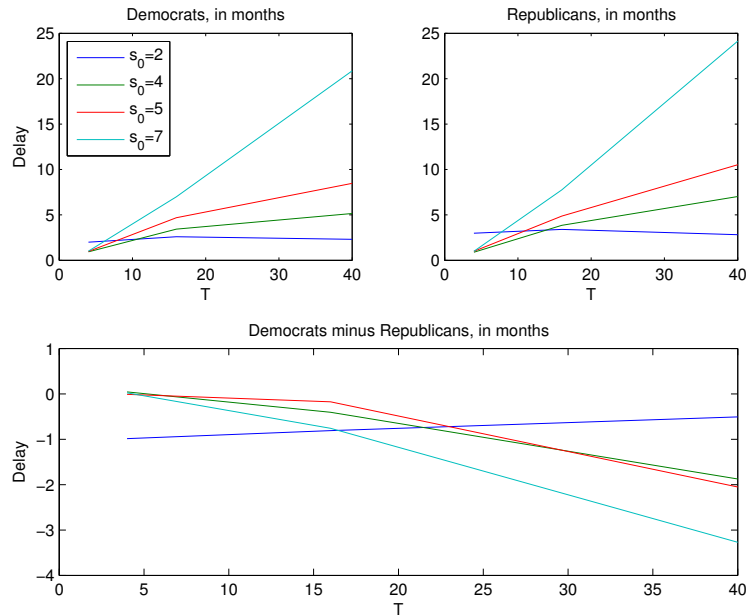
Figure 14
Average Real Wage Changes, Simulations
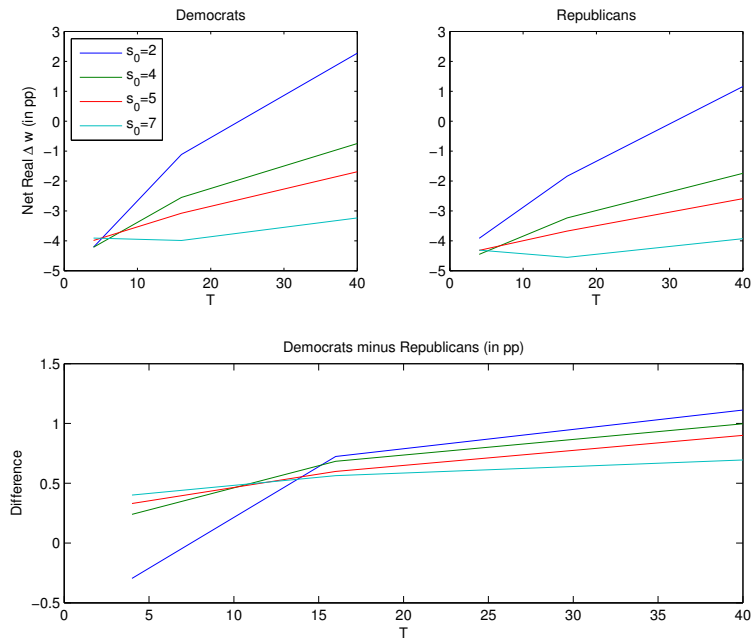
Figure 15
Average Real Pension Changes, Simulations
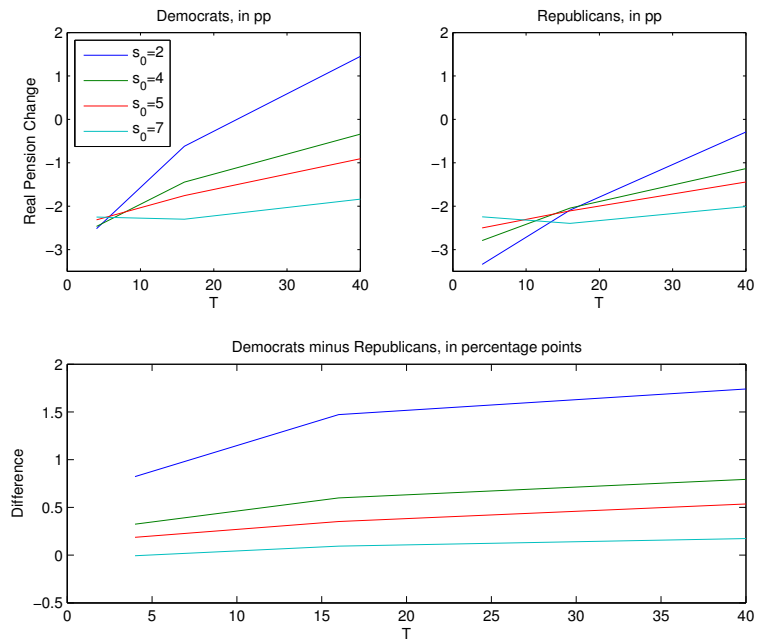
Looking at Figure 14, we see essentially the opposite dynamic, namely that wage growth is *more* sensitive to changes in T when unemployment is low. Again, the logic deals with the flow benefits and proximity to the disagreement costs. In the good state, when T is large, early in bargaining the governor has strong incentives to reach agreement in order to guarantee himself high flow benefits for the remaining periods. Thus, the governor will agree to relatively large compensation increases. However, when T is small, the potential number of periods of foregone flow benefits is small, so the governor is less willing to agree to compensation increases. Note that the union does not face a symmetric type of pressure to agree early, since it does not have associated direct flow benefits. Therefore, it disproportionately benefits from the increases in T in good states. In contrast, in bad states, flow benefits are low and thus there is little pressure on the governor to reach agreement in any value of T. Thus, for all values of T, agreements tend to feature relatively low wage growth. Note from Figure 15 that the dynamics of pension growth are very similar.

The counterfactuals also demonstrate that, in almost all specifications, Democrats agree to higher wage and pension growth than Republicans. For instance, when $T = 16$ and $s_0 = 4$, agreements with Democrats feature wage growth approximately 0.684 percentage points higher than those with Republicans, with pension growth approximately 0.600 percentage points higher. To put the differences in wage growth in monetary terms, suppose the average salary of public sector workers is \$50,000 per year.[34] Furthermore, in 2008 the median number of state public sector workers covered by a contract was 81,574. If each of these workers received an additional 0.684 percentage point raise in a given year, that would correspond to approximately \$27,891,943 in increased wage outlays. While not enormous in per capita terms, if this compounds over time, or if there are spillovers to public sector workers not covered by a contract, then the difference in public sector compensation may become significant.

In addition, I investigate the effect of pension funding levels on bargaining outcomes. I solve the model for a bargaining spell with an upcoming open seat election, and with

---

[34]This is approximately correct, computed from American Community Survey 2008 data.

median levels of all variables (including $s_0 = 4$ and $T = 16$) except for party and funding levels. I vary funding levels between the 10th, 50th, and 90th percentiles in the data and simulate the model $N = 1000$ times to understand the effect of initial pension funding on bargaining outcomes, depending on initial state and party. Table 4 shows the results for wage and pension changes from the simulations. The model implies a tradeoff between pensions and wages. A funding level of 106% versus 61% implies higher pension growth by about 1.13 percentage points for Democrats and 0.9 percentage points for Republicans, and similar sized declines in wage growth. Thus, in the model total compensation does not significantly change as funding increases. Rather, improvements in pension funding tend push the governor to substitute pensions in place of wages, since raising pensions is relatively less costly.

Table 4
SIMULATED MEANS, NET CHANGES (IN PP), BY FUNDING

	Democrat			Republican		
	61%	85%	106%	61%	85%	106%
Wages	-1.767	-2.396	-2.949	-2.566	-3.068	-3.532
Pensions	-1.939	-1.349	-0.806	-2.428	-1.942	-1.515

The final exercise I perform is the following. In order to determine the degree to which model outcomes are driven by the unobservable shock, I solve and simulate the model while shutting down γ_t. Overall, the model without shocks generates slightly more delay and disagreement, but still involves a significant probability of agreement. I find that the average delay with no unobservable shocks is 8.17, with a 7.3% probability of disagreement. This contrasts with an average delay of 7.07 in the full model, and a probability of disagreement of 2.6%. Furthermore, the mean wage and pension growth implied by the model with the unobservable shock, approximately -1.8 and -1.7 percentage points, respectively, are very similar to those in the model without the shock, -2.2 and -1.9 percentage points. Thus, the introduction of the unobservable component smooths the distributions of bargaining outcomes, but overall it is not the dominant driver of model outcomes.

1.7 Conclusion

Public sector compensation is a relatively sizable share of U.S. GDP, and has strained state budgets in recent years as tax revenues declined due to the Great Recession. Despite this, understanding the determinants of public sector compensation, and in particular the role of political variables, is typically less focal in the literature. To this end, I develop a stochastic bargaining model in which a governor and public union bargain over compensation. The political and economic variables in the model affect both the propensity of sides to reach agreement, as well as the terms of an agreement. For instance, in bad economic times, both sides may want to postpone agreement until times get better and, say, the governor has more political capital or the state budget is more flexible. Furthermore, the sensitivity to these factors may depend on, for example, political party. Democrats may be more amenable to reaching favorable terms of agreement to the union relative to Republicans. Thus, the model has implications for the relationship between political and economic variables and bargaining outcomes, namely delay, wage growth, and pension growth.

I use hand-collected data from various sources, including collective bargaining agreements for state-level public sector unions, to calibrate the model and investigate the effects of political and economic variables on bargaining outcomes. Using the calibrated model to run counterfactuals, I find that wage and pension growth outcomes are higher under Democratic governors, while delay is significantly shorter. Additionally, higher unemployment rates at the beginning of bargaining tend to have a negative impact on compensation levels. The magnitude of these responses is sensitive to time before the next election. I also find that bargaining spells for unions with more well-funded pension plans tend to generate higher pensions. Overall, the model sheds insight into how both political and economic factors can affect public sector compensation, a topic that has arguably been understudied in the past.

1.8 Appendix

1.8.1 Additional Tables

Table 5
SIMULATION PARAMETERS

Parameter		Value
Dem. gov. disagree cost, $s = 1$:	$d_{g,1}^D$	-18.552
Dem. gov. disagree cost, $s = S$:	$d_{g,S}^D$	-3.060
Rep. gov. disagree cost, $s = 1$:	$d_{g,1}^R$	-24.473
Rep. gov. disagree cost, $s = S$:	$d_{g,S}^R$	-4.951
Dem. flow benefit, $s = 1$:	n_1^D	0.950
Dem. flow benefit, $s = S$:	n_S^D	-0.650
Rep. flow benefit, $s = 1$:	n_1^R	0.900
Rep. flow benefit, $s = S$:	n_S^R	-0.800
Dem. one-time wage cost, $s = 1$:	$c_{1,1}^D$	69.063
Dem. one-time wage cost, $s = S$:	$c_{1,S}^D$	340.000
Rep. one-time wage cost, $s = 1$:	$c_{1,1}^R$	74.375
Rep. one-time wage cost, $s = S$:	$c_{1,S}^R$	425.000
Dem. one-time pension cost, $s = 1$:	$c_{2,1}^D$	105.625
Dem. one-time pension cost, $s = S$:	$c_{2,S}^D$	381.875
Rep. one-time pension cost, $s = 1$:	$c_{2,1}^R$	120.250
Rep. one-time pension cost, $s = S$:	$c_{2,S}^R$	461.500
One-time pension cost adjustment, funding:	c_2^f	40.000
Wage effect on vote share:	ϕ_1	31.330
Pension effect on vote share:	ϕ_2	14.725
Voter preference distribution parameter:	λ	0.040
CPVI effect on vote share:	χ_1	0.128
Open seat effect on vote share:	χ_2	-2.271
Shifter for vote share, $s = 1$:	μ_1	-17.304
Shifter for vote share, $s = S$:	μ_S	-18.753
Union disagree cost, $s = 1$:	$d_{u,1}$	2544.795
Union disagree cost, $s = S$:	$d_{u,S}$	5649.630
Average pension weight, non-PS unions:	$mean(\eta_j)$	0.428
Average pension weight, PS unions:	$mean(\eta_{ps,j})$	0.345
Weight on union terminal utility:	Υ	100
Common discount factor:	δ	0.990
Coefficient of relative risk aversion:	σ	3
$\gamma_{t,b,j}$-shock variance, $s = 1$:	θ_1	20
$\gamma_{t,j}$-shock variance:	θ_2	20
Baseline probability union proposes:	$\overline{\Pi}_u$	0.450
Proposal probability adjustment, MEM:	β_1	0.100

1.8.2 Proofs

Proof of Proposition 1. To see that a subgame perfect equilibrium (SPE) exists, first I
let $v_i(s, t, \kappa)$ denote the *ex-ante* value function for agent $i \in \{g, u\}$ in state $s \in \mathcal{S}$ and date
$t \in \{1, ..., T\}$ when κ is the proposer. For notational simplicity, I abstract from dependence
of the value functions on other observables. "Ex-ante" here refers to prior to the realization
of the γ_t shock, but after the state and proposer are realized.

Now, consider the case where the union is responding to a proposal $(\widehat{w}, \widehat{p})$ from the gover-
nor in period T and state s. For notational convenience, I let $CV_{s,T}^u \equiv u_u(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) +$
$\delta \mathbb{E}_{\kappa', s'|s} v_u(s', T+1, \kappa')$, which is the period T value to the union from not reaching agree-
ment. Clearly, in any SPE it must be that the union rejects an offer if $U_u(\widehat{w}, \widehat{p}, X, \overline{\tau}_w, \overline{\tau}_p) <$
$CV_{s,T}^u$ and accepts if $U_u(\widehat{w}, \widehat{p}, X, \overline{\tau}_w, \overline{\tau}_p) > CV_{s,T}^u$. Consider an offer that the union is indif-
ferent between accepting and rejecting, that is satisfying $U_u(\widehat{w}, \widehat{p}, X, \overline{\tau}_w, \overline{\tau}_p) = CV_{s,T}^u$. Let
$\overline{q}_{s,T}^u(\widehat{w}, \widehat{p})$ denote the union's probability of accepting in period T and state s. I consider
four different cases:

1. Suppose $(\widehat{w}, \widehat{p})$ solves the program:

$$\max_{(w,p) \in \mathbb{R}_+^2} U_g(w, p, s, T; fun, Y, D, \underline{w}, \underline{p}) + \gamma_T \tag{10}$$

$$\text{s.t. } U_u(w, p, X, \overline{\tau}_w, \overline{\tau}_p) = u_u(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) + \delta \mathbb{E}_{\kappa', s'|s} v_u(s', T+1, \kappa'),$$

   and $U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) + \gamma_T > o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', T+1, \kappa')$. For con-
   ciseness, I will assume that $(\widehat{w}, \widehat{p})$ is the unique solution to the program.[35] Note that
   in this case, the governor would prefer to reach agreement under $(\widehat{w}, \widehat{p})$ than to pass.
   If $\overline{q}_{s,T}^u(\widehat{w}, \widehat{p}) < 1$, then there would be no optimal offer for the governor. To see this,

---

[35]If there are multiple maximizers, then at least one of the solutions must satisfy $\overline{q}_{s,T}^u(w, p) = 1$ by the
same logic. Note, however, that no matter which solution satisfies $\overline{q}_{s,T}^u(w, p) = 1$ the payoffs are identical
since all "maximizers" must give the same payoff to the governor and union by definition.

the governor's utility from offering $(\widehat{w}, \widehat{p})$ would be:

$$\overline{q}^u_{s,T}(\widehat{w}, \widehat{p})(U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) + \gamma_T) + (1 - \overline{q}^u_{s,T}(\widehat{w}, \widehat{p}))(o(s, D) +$$
$$\delta\mathbb{E}_{\kappa', s'|s}v_g(s', T+1, \kappa')) < U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) + \gamma_T.$$

Since union utility is increasing in wages, for all $\varepsilon > 0$, $\overline{q}^u_{s,T}(\widehat{w} + \varepsilon, \widehat{p}) = 1$. Thus, there exists $\overline{\varepsilon}$ such that for all $\varepsilon \in (0, \overline{\varepsilon})$

$$\overline{q}^u_{s,T}(\widehat{w}, \widehat{p})(U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) + \gamma_T) + (1 - \overline{q}^u_{s,T}(\widehat{w}, \widehat{p}))(o(s, D) +$$
$$\delta\mathbb{E}_{\kappa', s'|s}v_g(s', T+1, \kappa')) < U_g(\widehat{w} + \varepsilon, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) + \gamma_T.$$

Furthermore, for any alternative offer (w', p') that is accepted with positive probability, there exists $\varepsilon' < \overline{\varepsilon}$ such that:

$$U_g(\widehat{w} + \varepsilon', \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) > U_g(w', p', s, T; fun, Y, D, \underline{w}, \underline{p}).$$

This is because any other offer satisfying $U_u(w', p', s, T; fun, Y, D, \underline{w}, \underline{p}) \geq CV^u_{s,T}$ must be such that:

$$U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) > U_g(w', p', s, T; fun, Y, D, \underline{w}, \underline{p}),$$

since $(\widehat{w}, \widehat{p})$ is the maximizer of the governor's utility subject to giving the union his continuation value.

Since $U_g(\widehat{w} + \varepsilon', \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) \to U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p})$ as $\varepsilon' \to 0$, this implies that for any alternative offer (w', p') that may be accepted, there exists a dominant one in $(\widehat{w} + \varepsilon', \widehat{p})$ for sufficiently small ε'. Thus, $\overline{q}^u_{s,T}(\widehat{w}, \widehat{p}) < 1$ would imply there is no optimal offer, since the governor would prefer to make ε as small as possible, but still strictly positive. Therefore, in any SPE it must be that $\overline{q}^u_{s,T}(\widehat{w}, \widehat{p}) = 1$ if $(\widehat{w}, \widehat{p})$ solves program (10) and $U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) > o(s, D) + \delta\mathbb{E}_{\kappa', s'|s}v_g(s', T +$

$1, \kappa')$. In turn, the governor would strictly prefer to offer $(\widehat{w}, \widehat{p})$ than to make an alternative offer or to pass. Therefore, in such cases, there are unique payoffs at date T and state s, namely $U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p})$ for the governor and $CV^u_{s,T}$ for the union.

2. Next, suppose $(\widehat{w}, \widehat{p})$ solves (10), but instead $U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) < o(s, D) + \delta\mathbb{E}_{\kappa', s'|s}v_g(s', T+1, \kappa')$. Then any $\overline{q}^u_{s,T}(\widehat{w}, \widehat{p}) \in [0, 1]$ is possible in an SPE. This is because the union is indifferent between accepting and rejecting, and furthermore the governor would strictly prefer to pass instead of offering $(\widehat{w}, \widehat{p})$, unless $\overline{q}^u_{s,T}(\widehat{w}, \widehat{p}) = 0$ in which case he is indifferent. Thus, even though in this situation there are multiple possible strategies in equilibrium, all of them yield the same payoffs to players. Therefore, again there are unique payoffs in these cases at state s and date T, namely each player's continuation value.

3. Next, suppose $(\widehat{w}, \widehat{p})$ does not solve (10). If the solution to (10), denoted $(\widehat{w}', \widehat{p}')$, satisfies $U_g(\widehat{w}', \widehat{p}', s, T; fun, Y, D, \underline{w}, \underline{p}) > o(s, D) + \delta\mathbb{E}_{\kappa', s'|s}v_g(s', T+1, \kappa')$, then from part (1) it must be that $q^u_{s,T}(\widehat{w}', \widehat{p}') = 1$. Then, since the governor would strictly prefer $(\widehat{w}', \widehat{p}')$ to offer $(\widehat{w}, \widehat{p})$, any $q^u_{s,T}(\widehat{w}, \widehat{p}) \in [0, 1]$ is possible in equilibrium. If instead $U_g(\widehat{w}', \widehat{p}', s, T; fun, Y, D, \underline{w}, \underline{p}) \leq o(s, D) + \delta\mathbb{E}_{\kappa', s'|s}v_g(s', T+1, \kappa')$, it must be $U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) < o(s, D) + \delta\mathbb{E}_{\kappa', s'|s}v_g(s', T+1, \kappa')$, in which case the governor would always prefer passing to offering $(\widehat{w}, \widehat{p})$ regardless of $q^u_{s,T}(\widehat{w}, \widehat{p})$, again implying that any $q^u_{s,T}(\widehat{w}, \widehat{p}) \in [0, 1]$ is possible. However, note that in this case, $(\widehat{w}, \widehat{p})$ will not be on the path of play since there is a strictly preferred strategy for the governor, given the equilibrium strategies of the union. Again, this does not affect the uniqueness of payoffs.

4. Finally, suppose that $(\widehat{w}, \widehat{p})$ solves (10) and $U_g(\widehat{w}, \widehat{p}, s, T; fun, Y, D, \underline{w}, \underline{p}) = o(s, D) + \delta\mathbb{E}_{\kappa', s'|s}v_g(s', T+1, \kappa')$. In this case, both sides are indifferent between the best possible agreement and passing. Therefore, any $\overline{q}^u_{s,T}(\widehat{w}, \widehat{p}) \in [0, 1]$ is possible, and the governor will be indifferent between passing and offering $(\widehat{w}, \widehat{p})$. However, in any of these cases,

the equilibrium payoffs will be the same: $o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', T + 1, \kappa')$ for the governor and $CV^u_{s,T}$ for the union.

Identical logic applies when the union is the proposer. Therefore, in period T, as long as there exists an optimal offer $(\widehat{w}, \widehat{p})$ or all offers will be rejected – which I show is the case below – since the continuation payoffs are exogenous, the SPE payoffs in each state s are unique. Given that the SPE payoffs at T are unique, the continuation values at time $T - 1$ are also unique, so the same logic applies to this period. Working backwards implies the entire path of SPE payoffs are unique.

To complete the proof, I argue that either: (i) there always exists an optimal offer at each s and t and for each proposer, or (ii) if an optimal offer doesn't exist, it implies that passing or making an offer that will be rejected with certainty is the optimal action for the proposer. Letting:

$$A \equiv \frac{\left(\frac{1 - \delta^{T-t+1}}{1 - \delta}\right) + \delta^{T-t+1} \Upsilon}{1 - \sigma},$$

the constraint when the governor determines its optimal offer is:

$$A \left((w(1 - \overline{\tau}_w))^{1-\sigma} + \eta(X) \left(p(1 - \overline{\tau}_p) \right)^{1-\sigma} \right) = CV^u_{s,t}.$$

Algebraic manipulation yields

$$p^g_{s,t}(w) = \frac{1}{1 - \overline{\tau}_p} \left(\frac{1}{\eta(X)} \left(\frac{CV^u_{s,t}}{A} - (w(1 - \overline{\tau}_w))^{1-\sigma} \right) \right)^{\frac{1}{1-\sigma}}, \qquad (11)$$

where $p^g_{s,t}(w)$ is the pension level, given a value of w that satisfies the constraint. Note that for $\sigma > 1$,[36] A is negative. Furthermore, so long as the disagreement costs for the union are not too negative, the continuation value $CV^u_{s,t}$ is also negative, making $\frac{CV^u_{s,t}}{A} > 0$. I consider the case where $CV^u_{s,t} > 0$ below.

---

[36]Certainly σ could be less than one, and similar proofs would apply. To ease exposition, the proofs are written assuming $\sigma > 1$ since the calibrated coefficient of relative risk aversion is larger than one.

Also note that there is a lower bound on w, denoted w_{LB} that satisfies

$$0 = \frac{CV_{s,t}^u}{A} - (w_{LB}(1 - \overline{\tau}_w))^{1-\sigma} \iff$$

$$w_{LB} = \frac{1}{1 - \overline{\tau}_w} \left(\frac{CV_{s,t}^u}{A} \right)^{\frac{1}{1-\sigma}}.$$

This is because for all $w \leq w_{LB}$, there is no finite pension level that induces acceptance on the part of the union.

Then, the problem of the governor is:

$$\max_{w \geq 0} U_g(w, p_{s,t}^g(w), s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma_t$$

The FOC of this is given by:

$$-c_1(s, D) - (c_2(s, D) - c_2^f fun) \frac{\partial p_{s,t}^g(w^*)}{\partial w} - \delta^{T-t+1} \mathbb{E}_{s_{T+1}|s} \Big[f(\phi_1(w^* - \underline{w}) + \quad (12)$$

$$\phi_2(p_{s,t}^g(w^*) - \underline{p}) - \mu(s_{T+1}, Y)) \left(\phi_1 + \phi_2 \frac{\partial p_{s,t}^g(w^*)}{\partial w} \right) \Big] = 0,$$

where w^* is the optimal wage offer and:

$$\frac{\partial p_{s,t}^g(w)}{\partial w} = -\frac{1 - \overline{\tau}_w}{1 - \overline{\tau}_p} \left(\frac{1}{\eta} \right)^{\frac{1}{1-\sigma}} \left(\frac{CV_{s,t}^u}{A} - (w(1 - \overline{\tau}_w))^{1-\sigma} \right)^{\frac{\sigma}{1-\sigma}} (w(1 - \overline{\tau}_w))^{-\sigma},$$

which is strictly negative so long as $w > w_{LB}$, which it must be at the optimum.

As $w \to w_{LB}$ from above, $\frac{\partial p_{s,t}^g(w)}{\partial w} \to -\infty$, and therefore the LHS of (12) approaches ∞. Also, as $w \to \infty$, since $f(\cdot)$ is bounded and $\frac{\partial p_{s,t}^g(w^*)}{\partial w} \to 0$, the LHS of (12) approaches a strictly negative number. Thus, there exists at least one wage that satisfies (12). Furthermore, for at least one of these wages, the second order condition is satisfied since the LHS of (12) must be decreasing at one or more of them. Therefore, a finite optimal offer for the governor exists when $CV_{s,t}^u < 0$. Even if multiple w^* maximize the governor's utility, by definition they yield the same utility to both the governor and the union.

If instead $CV^u_{s,t} \geq 0$ when $\sigma > 1$, this implies that the continuation value for the union is so large that there is no positive and finite level of w and p the union is willing to accept. Intuitively, when $\sigma > 1$, flow utilities are always negative for any positive w and p. Therefore, if the continuation value for the union is positive, there is no way to satisfy the union. In this case, any offer will be rejected, which implies unique utilities for both the governor and the union.

Now consider the case in which the union proposes. The union solves for the optimal contract conditional on inducing acceptance:

$$\max_{(w,p)\in\mathbb{R}^2_+} U_u(w,p,X,\overline{\tau}_w,\overline{\tau}_p) \tag{13}$$

$$\text{s.t. } U_g(w,p,s,t;fun,Y,D,\underline{w},\underline{p}) + \gamma_t = o(s,D) + \delta\mathbb{E}_{\kappa',s'|s}v_g(s',t+1,\kappa').$$

Similar to the case with the governor, for each possible wage w, one can solve for the pension level that satisfies the constraint. However, in this case, the pension level also depends on γ_t. Thus, the function $p^u_{s,t}(w,\gamma_t)$ is defined to be the pension level, given w and γ_t that satisfies the constraint. Note that this cannot be solved for in closed-form, though it can be solved for numerically.

Plugging this into the objective function, the FOC for the modified problem is:

$$(w^*)^{-\sigma}(1-\overline{\tau}_w)^{1-\sigma} + \eta(X)p^u_{s,t}(w^*,\gamma_t)^{-\sigma}(1-\overline{\tau}_p)^{1-\sigma}\left(\frac{\partial p^u_{s,t}(w^*,\gamma_t)}{\partial w}\right) = 0, \tag{14}$$

where again w^* is the optimal wage offer and the implicit function theorem from the constraint of (13) gives:

$$\frac{\partial p^u_{s,t}(w,\gamma_t)}{\partial w} = -\frac{c_1(s,D) + \delta^{T-t+1}\mathbb{E}f(\phi_1(w^*-\underline{w}) + \phi_2(p^u_{s,t}(w^*,\gamma_t)-\underline{p}) - \mu(s_{T+1},Y))\phi_1}{c_2(s,D) + \delta^{T-t+1}\mathbb{E}f(\phi_1(w^*-\underline{w}) + \phi_2(p^u_{s,t}(w^*,\gamma_t)-\underline{p}) - \mu(s_{T+1},Y)\phi_2} < 0.$$

First, suppose $o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa')$ is sufficiently small such that:

$$U_g(0, 0, s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma_t > o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa'). \qquad (15)$$

If this inequality is satisfied, it implies that at a wage-pension offer of $(w, p) = (0, 0)$, the governor receives enough utility to strictly prefer the contract to passing. I define w_{UB} to be the wage such that $p_{s,t}^u(w_{UB}, \gamma_t) = 0$, which is defined by:

$$U_g(w_{UB}, 0, s, t; fun, Y, D, \underline{w}, \underline{p}) = o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa').$$

Since the LHS is decreasing in the wage and approaches $-\infty$, as long as (15) is satisfied, intermediate value theorem implies $w_{UB} > 0$ exists. As $w \to w_{UB}$, since $\frac{\partial p_{s,t}^u(w, \gamma_t)}{\partial w} < 0$ the LHS of (14) approaches $-\infty$. Furthermore, as $w \to 0$, the LHS of (14) approaches ∞. Thus, by the same logic as in the governor's proposal, there must exist an optimal offer.

If alternatively (15) is not satisfied, this implies that the governor's continuation value is so high that he would (weakly) prefer to reject even an offer of $w = p = 0$. Thus, in this case, it is also clearly optimal for the union to pass, or make an offer that will certainly be rejected. Thus, at each date, state, and γ_t, there is always an optimal offer (or decision to pass) for the proposer, which completes the proof. $\qquad \square$

Proof of Lemma 1. To begin the proof, I show that an optimal offer exists if and only if $\gamma > \gamma^*$. To see this, given a value of γ, an optimal offer *does not* exist if and only if for all (w, p):

$$U_g(w, p, s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma < o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa').$$

Since U_g is decreasing in w and p, if this inequality holds for γ, then it must also hold for all $\gamma' < \gamma$. Thus, with γ^* satisfying:

$$U_g(0, 0, s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma^* = o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa'),$$

for all $\gamma > \gamma^*$ an optimal contract exists, and for all $\gamma < \gamma^*$, one does not exist. Thus, for all $\gamma < \gamma^*$, there will be no agreement.

Next, consider $\gamma > \gamma^*$. Let $(w_{s,t}^u(\gamma_t), p_{s,t}^u(\gamma_t))$ denote the optimal offer when the union proposes and the shock is γ_t. I now show that the utility of the union must be increasing in γ_t. To see this, consider two realizations of the shock, $\gamma' > \gamma$. Then, consider the optimal wage and pension when γ is the realization, denoted (w^*, p^*). Under γ':

$$U_g(w^*, p^*, s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma' > U_g(w^*, p^*, s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma$$
$$= o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa'),$$

so (w^*, p^*) is still feasible when γ' is realized. Furthermore, since U_g is decreasing in w, there exists a $w' > w^*$ such that:

$$U_g(w', p^*, s, t; fun, Y, D, \underline{w}, \underline{p}) + \gamma' = o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa').$$

Furthermore since U_u is increasing in w, $U_u(w', p^*, X, \overline{\tau}_w, \overline{\tau}_p) > U_u(w^*, p^*, X, \overline{\tau}_w, \overline{\tau}_p)$. Therefore, the optimal wage and pension offer under γ' must provide strictly higher utility to the union than the optimal offer under γ.

Since utility of the union evaluated at the optimal offer is strictly increasing in γ, and since the union makes an offer if and only if:

$$U_u(w_{s,t}^u(\gamma_t), p_{s,t}^u(\gamma_t), s, t; fun, Y, D, \underline{w}, \underline{p}) \geq u_u(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) + \delta \mathbb{E}_{\kappa', s'|s} v_u(s', t+1, \kappa')$$

where the RHS does not depend on γ, the union follows a cutoff rule for making a proposal. The cutoff shock, denoted $\underline{\gamma}_{s,t}^u$, satisfies:

$$U_u(w_{s,t}^u(\underline{\gamma}_{s,t}^u), p_{s,t}^u(\underline{\gamma}_{s,t}^u), X, \overline{\tau}_w, \overline{\tau}_p) = u_u(\underline{w}, \underline{p}, X, \overline{\tau}_w, \overline{\tau}_p) + \delta \mathbb{E}_{\kappa', s'|s} v_u(s', t+1, \kappa').^{37}$$

This completes the proof. $\qquad\square$

Proof of Lemma 2. I prove by contradiction. Suppose instead that $\underline{\gamma}^u_{s,t} \neq \underline{\gamma}^g_{s,t}$. Suppose without loss of generality that $\underline{\gamma}^u_{s,t} > \underline{\gamma}^g_{s,t}$. By definition, it must be:

$$U_u(w^u_{s,t}(\underline{\gamma}^u_{s,t}), p^u_{s,t}(\underline{\gamma}^u_{s,t}), X, \bar{\tau}_w, \bar{\tau}_p) = u_u(\underline{w}, \underline{p}, X, \bar{\tau}_w, \bar{\tau}_p) + \delta \mathbb{E}_{\kappa', s'|s} v_u(s', t+1, \kappa')$$

$$U_g(w^u_{s,t}(\underline{\gamma}^u_{s,t}), p^u_{s,t}(\underline{\gamma}^u_{s,t}), s, t; fun, Y, D, \underline{w}, \underline{p}) + \underline{\gamma}^u_{s,t} = o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa')$$

$$U_u(w^g_{s,t}, p^g_{s,t}, X, \bar{\tau}_w, \bar{\tau}_p) = u_u(\underline{w}, \underline{p}, X, \bar{\tau}_w, \bar{\tau}_p) + \delta \mathbb{E}_{\kappa', s'|s} v_u(s', t+1, \kappa')$$

$$U_g(w^g_{s,t}, p^g_{s,t}, s, t; fun, Y, D, \underline{w}, \underline{p}) + \underline{\gamma}^g_{s,t} = o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa').$$

Note that since κ' is independent of the current proposer, the continuation values for each player are also independent of proposer. Therefore:

$$U_g(w^u_{s,t}(\underline{\gamma}^u_{s,t}), p^u_{s,t}(\underline{\gamma}^u_{s,t}), s, t; fun, Y, D, \underline{w}, \underline{p}) + \underline{\gamma}^u_{s,t} = U_g(w^g_{s,t}, p^g_{s,t}, s, t; fun, Y, D, \underline{w}, \underline{p}) + \underline{\gamma}^g_{s,t},$$

which holds if and only if:

$$0 < \underline{\gamma}^u_{s,t} - \underline{\gamma}^g_{s,t}$$
$$= U_g(w^g_{s,t}, p^g_{s,t}, s, t; fun, Y, D, \underline{w}, \underline{p}) - U_g(w^u_{s,t}(\underline{\gamma}^u_{s,t}), p^u_{s,t}(\underline{\gamma}^u_{s,t}), s, t; fun, Y, D, \underline{w}, \underline{p}).$$

Finally, this holds if and only if:

$$U_g(w^g_{s,t}, p^g_{s,t}, s, t; fun, Y, D, \underline{w}, \underline{p}) > U_g(w^u_{s,t}(\underline{\gamma}^u_{s,t}), p^u_{s,t}(\underline{\gamma}^u_{s,t}), s, t; fun, Y, D, \underline{w}, \underline{p}).$$

However, this also implies:

$$U_g(w^g_{s,t}, p^g_{s,t}, s, t; fun, Y, D, \underline{w}, \underline{p}) + \underline{\gamma}^u_{s,t} > o(s, D) + \delta \mathbb{E}_{\kappa', s'|s} v_g(s', t+1, \kappa').$$

which in turn implies that the union's offer is suboptimal. To see this, note that the union is indifferent between $(w^g_{s,t}, p^g_{s,t})$ and $(w^u_{s,t}(\underline{\gamma}^u_{s,t}), p^u_{s,t}(\underline{\gamma}^u_{s,t}))$. Furthermore, since the governor's utility is decreasing in the wage and the union's is increasing, there exists an offer $(w', p^g_{s,t})$ with $w' = w^g_{s,t} + \epsilon$ that the union strictly prefers to $(w^g_{s,t}, p^g_{s,t})$ and $(w^u_{s,t}(\underline{\gamma}^u_{s,t}), p^u_{s,t}(\underline{\gamma}^u_{s,t}))$ such

that:

$$U_g(w', p^g_{s,t}, s, t; fun, Y, D, \underline{w}, \underline{p}) + \underline{\gamma}^u_{s,t} = o(s, D) + \delta\mathbb{E}_{\kappa',s'|s}v_g(s', t+1, \kappa').$$

Thus, the offer $(w^u_{s,t}(\underline{\gamma}^u_{s,t}), p^u_{s,t}(\underline{\gamma}^u_{s,t}))$ is suboptimal, generating a contradiction. $\qquad\qquad\square$

1.8.3 Calculation of Tax Rates and Pensions

In order to calculate pension levels, I follow closely a procedure from Beshears et al (2011), which makes assumptions about a representative worker's age, income stream, tenure in the public sector, and marital status.[38] Then, using information on public sector pension retirement factors and years used to compute final average salary for the particular union, I can compute the gross pension for the representative worker. Given this I can compute a gross pension for the initial year of retirement. I then use information from Beshears et al (2011) on the tax treatment of public pensions in each geographic state, as well as NBER's TAXSIM program, to compute the average tax rate on retirement income, which in turn gives the after-tax pension for the representative worker.

More specifically, I assume that the "representative" worker has the following characteristics:

- At age 40, the worker has a $50,000 salary;

- The worker expects on average a 1% real wage increase for the remainder of her career;

- The worker will retire at age 65;

- The worker is single;

- By retirement, the worker will have been in the public sector for 30 years.

Let the assumed salary for the representative worker at age a be denoted Sal_a. Then, given that the number of years used to compute final average salary is y and the retirement factor

---

[38]Marital Status is only relevant through its impact on the tax rate the worker will face.

is RF, the pension (denoted Pen) for the worker is

$$Pen = \left(\frac{1}{y} \sum_{a=0}^{y-1} Sal_{65-a} \right) \times RF \times 30,$$

where 30 comes from the assumed 30 years of tenure in the public sector.

The level of the initial pension is determined from the initial salary flow and initial pension parameters. To compute p^*, one needs an assumption on the *new* income stream upon agreement. Suppose the agreed upon gross wage increase is $(1 + g)$. I assume that the worker continues to expect a 1% real wage increase (on average) for the remainder of her career. Thus, if the initial age a expected salary is $\overline{Sal}_a$, then the new expected salary will be $Sal_a^* = \overline{Sal}_a(1 + g)$. In other words, the income stream will shift up by $(1 + g)$. Then, the p^* computations are based on this new salary stream *and* the new pension parameters. I normalize all pensions and wages by the initial wage level, i.e. \$50,000 by assumption. Thus, $\underline{w} \equiv 1$, w^* is the gross real wage increase, and $\underline{p}$ and p^* are the initial and agreed upon pensions, respectively, as a percentage of the initial salary (\$50,000).

To determine $\overline{\tau}_p$, I draw from Beshears et al (2011), which provides the tax treatment of public sector pensions. For instance, in some states, the full pension is taxable, while in other states part or all of a public pension is not taxable.[39] I use TAXSIM to obtain the average federal and FICA tax rates using the initial pension level, Pen, as taxable income. Furthermore, letting ρ be the fraction of public pension income that is taxable at the state level, I use TAXSIM to obtain the average state tax rate using ρPen as taxable income. These combined yield $\overline{\tau}_p$. For simplicity, I assume that this average tax rate also applies for any new agreed upon pension.[40] To obtain $\overline{\tau}_w$, given the assumption on the "representative" worker, I simply take the assumed salary of \$50,000 and compute the state, federal, and FICA tax rates.

---

[39] At the federal level, public sector pensions are fully taxable in all states.

[40] Due to progressive taxation, this is an approximation, not an exact calculation. Technically, one would use the *marginal* tax rates. However, since most changes are relatively small, the average taxes generally provide a very close approximation.

Note that the assumption on the level of initial annual wage ($50,000) is not particularly important. Alternative assumptions for the initial wage would only have a meaningful effect on outcomes through its effects on the average tax rate, since everything is normalized by the initial wage.

Chapter 2 : Campaign Spending and Strategy in U.S. Congressional Elections[*]

2.1 Introduction

Candidates running for political office spend a vast and ever-growing sum of money. In the 2012 cycle, candidates running for seats in the U.S. Congress spent about \$1.9 billion, representing an increase of almost 50% in real terms relative to 2000. Despite this significant sum of funds channeled to political campaigns, there seems to be no consensus among social scientists as to the impact of this money on political outcomes.[41] For example, Feldman and Jondrow (1984), Ragsdale and Cook (1987), and Levitt (1994) find no statistically significant effect of incumbent spending on outcomes - and perhaps even a negative effect - whereas Abramowitz (1988), Grier (1989), Moon (2006), and da Silveira and de Mello (2011) find a positive and statistically significant effect. Furthermore, most of this literature is unable to capture the heterogeneity of campaign spending effects across candidates. In this chapter, I propose a new empirical framework that explicitly models the heterogeneity in the use and effect of campaign funds. To this end, I use a model of campaigning that allows funds to be

---

[*]This chapter is co-authored with Ekim Cem Muyan.

[41]See, for example, Stratmann (2005)

spent on different campaign strategies, which may affect election outcomes differentially. I argue that an understanding of the impact of campaign funds on elections is possible only when the heterogeneous effects of campaigning strategies is uncovered.

This approach enables us to investigate campaigning strategies employed by the candidates running for political office. In particular, I model and analyze the campaign tone (positivity or negativity) a candidate uses. In fact, understanding campaign strategies is of interest in and of itself. Evidence suggests that campaigns have become increasingly negative in tone since 2000. For instance, Fowler and Ridout (2013) point out that in the 2000 presidential election, approximately 60% of ads were negative.[42] In 2012, approximately 85% of the total ads were negative. This rapid increase in negativity has sparked wide and often critical commentary of such advertisements.[43] The particular channel I investigate is built on the anecdotal and empirical evidence that suggests negative advertising may discourage voter turnout. For instance, it is widely believed that heavy negative campaigning between the two major party candidates in the 2000 Minnesota Gubernatorial elections depressed their turnout, which opened the door for the third-party candidate Jesse Ventura to win the election.[44] There is also some concern that negative campaigning may contribute to polarization or voter fatigue.[45] This feature of campaign strategy is often overlooked in the empirical literature.[46] The political science literature has often found that not only does negative advertising differ from positive advertising in its overall effects on voters, but also the effects vary *across* different groups of potential voters.[47] For instance, negative advertising may have a demobilizing effect on ideological voters, while positive advertising may

---

[42]Negativity is measured as attack or contrasting advertising, which is typical in the political science literature.

[43]See, for instance, *Washington Post*, Feb. 20, 2012, "Study: Negative Campaign Ads Much More Frequent, Vicious Than In Primaries Past."

[44]See, for example, Lentz (2001).

[45]See, for instance, Ansolabehere, Iyengar, Simon, and Valentino (1994), which in an experimental setting found evidence that negative advertisements "weakened confidence in the responsiveness of electoral institutions and public officials. As campaigns become more negative and cynical, so does the electorate."

[46]Often, any notion of strategy besides overall spending is overlooked. Three important exceptions are Stromberg (2008), which looks at presidential campaign stops, Nalebuff and Shachar (1999), which investigates the exertion of candidates' effort to increase participation, and Gordon and Hartmann (2013a), which analyzes the optimal allocation of advertising across states under the Electoral College. Even among these cases, none includes a choice of overall negativity of the campaign.

[47]See, for example, Ansolabehere and Iyengar (1997), and several others discussed in section 2.4.1.

be more effective in attracting swing voters. To the extent that optimal campaign strategies differ systematically across different types of candidates and elections, it is important to understand the differential impact negativity can have on voting outcomes.

This chapter builds on the argument that campaign finance and strategies are heavily interrelated, and therefore should be analyzed together. To understand the true impact of a dollar on election outcomes, one must understand how that dollar will be spent in a campaign. This is because elections differ in terms of their fundamentals, which ultimately determine how effective the campaign strategies employed will be. Thus, the effect of campaign funds depends on the effectiveness of the strategy that will be used in equilibrium, which will vary across candidates and elections. Unless one understands *how* funds will be allocated among different strategies, one cannot be able to uncover their true impact on the outcomes of elections. In addition, the campaign strategies chosen depend on the available funds to the candidate and to his opponent. For example, a candidate might be more likely to engage in negative campaigning when both he and his opponent have large budgets, but may tend to be more positive when he has a large money advantage. Hence, approaching these two questions in isolation could result in misleading answers. The theoretical and empirical strategy tries to avoid this by focusing on this very important interplay of campaign funds and strategies.

To this end, I develop a model featuring a game between candidates that decide their campaign strategies. In particular, candidates decide on how to allocate their total budgets between positive and negative campaigning. I denote a candidate's campaigning that includes information only about himself as a positive one. On the other hand, when the campaigning includes information about the opponent, it is a negative one.[48] Each constituency has three types of voters: the base (ideological) voters for each candidate and swing voters. I assume that negative campaigning is a demobilizing tool: it demobilizes the supporters of the opponent at the expense of possibly alienating some of the candidate's own. On the other hand, positive campaigning is used to attract swing voters to vote in

---

[48]This categorization is the norm in the literature when measuring negative advertising.

favor of the candidate. More precisely, I assume that ideological voters decide only on their turnout. When they do, they vote for the specific candidate they support. Swing voters, on the other hand, always turnout to vote, but decide on whom to vote based on the (positive) campaigns of the candidates. Hence, in the model, a candidate campaigns negatively to reduce the turnout of the opponent's supporters at the cost of decreasing his own. Positive campaigning, on the other hand, increase the portion of swing voters who vote for her.

Elections differ from each other in the measure of voter groups. The measure of ideological voters for each candidate, and hence swing voters, vary across elections, which result in different equilibrium campaigning strategies for each election. In the empirical specification, these levels of support are drawn from a distribution depending on the election-specific observables. These draws are observed by the candidates while they are unobservable to the econometrician. Given initial support, candidates choose their allocations simultaneously, after which the election takes place and the winners are realized.

To infer the overall campaign strategy of the candidates, I use data from from the Wisconsin Political Advertisement project that records each television advertisement aired by a candidate. This dataset also records the tone of the advertisement. Hence, I assume that the TV campaigning strategy is representative of the overall strategy of the player. Moreover, to estimate the distribution of voters, I use data from the American National Election Study. I calibrate the model to match patterns of campaign tone observed in the data, and then use the calibrated model to understand the effects of spending and strategies on voting outcomes.

To see why a model is necessary to understand the impact of campaign funds and strategies, consider the following. Negativity may be a useful strategy for candidates who are trailing, as it may lead the base of supporters of a front-runner to shrink. Conversely, positive campaigning may be relatively more effective for a front-runner. This is in line with the observation in the data that incumbents tend to be much more positive than challengers. However, if one uses ex-post vote measures to try and infer the effectiveness of advertising,

one would tend to see negativity correlated with low vote shares and thereby conclude negative advertising is ineffective. Note that this goes beyond purely incumbent-challenger races, as even in open seat elections, ex-post vote margin is negatively correlated with campaigning negativity. Controlling for the endogenous decision of campaign negativity with respect to things like initial voter support is important in understanding the overall effectiveness of different campaign strategies.

The calibrated model suggests that campaign spending is relatively ineffective at increasing vote shares. For the average election, which has budgets of about $2.4 million, a 10% increase in one candidate's budget increases his expected vote differential by about 0.4 percentage points. This is roughly in line with results from Levitt (1994), among others. To understand the differential impact of campaign spending, I perform multiple exercises. For example, consider an election where the candidates have the same measure of supporters. When these candidates have no funds to campaign, they are expected to tie in the election. Now consider providing one candidate with a $2.1 million budget while the other $700,000 (which are approximately the 75th and 25th percentile of observed budgets, respectively). This yields a 2.5 percentage point improvement in the expected vote of the first candidate. While not insignificant in absolute terms, a $1.4 million advantage is approximately the same size as the median budget. Thus, 2.5 percentage points is arguably a relatively minor increases for such a sizable budget advantage. I employ other calculations to find that, albeit small, trailing candidates benefit from extra funds more than the leading ones. The model also implies that negative campaigning is relatively effective for candidates who face an opponent with a high level of initial support, while positive campaigning is relatively effective for candidates in elections where neither side has a particularly high initial support. Still, the differences in effectiveness are not large, especially given that overall effectiveness is low. Finally, the model implies slightly decreasing returns to spending. Both this feature and the relative effectiveness of negative campaigning for trailing candidates may explain why the previous literature tends to find challenger spending is relatively more effective than incumbent spending.

The rest of the chapter is organized as follows. Section 2.2 discusses the related literature while section 2.3 focuses on the underlying institutional framework. Section 2.4 describes the model and the proof of the existence of a unique equilibrium. Section 2.5 discusses the data, and section 2.6 describes the calibration of the model. Finally, section 2.7 provides the description and results from model simulations, and section 2.8 concludes.

2.2   Related Literature

The key feature of the model is the candidate's decision of how to allocate his budget between positive and negative campaigning. Several theoretical papers focus on this decision. One of the earliest examples is Harrington and Hess (1996), which studies negative campaigning in a spatial framework. In their model, the negativity of a campaign depends primarily on the personal attributes of the candidate. Later works focus on the signaling game associated with advertising when candidate qualifications are unknown to voters (see Bhattacharya (2012) and Hao and Li (2013), for instance).

For tractability, I abstract from the spatial framework and the signaling aspect of political advertising and focus on the direct effect by using an "influence function" (Bhattacharya 2012) that affects voter support for each candidate. The theoretical framework I utilize for the campaign stage of the model is similar to Skaperdas and Grofman (1995). They model a two-candidate competition where each candidate decides on positive and negative campaigning levels given fixed and equal budgets. Negative campaigning by a candidate depresses turnout (for both his own and his opponent's supporters) while positive campaigning influences undecided voters. Through this setup, they argue that they can broadly match some regularities of political competition - namely, that the front-runner chooses more positive advertising than his opponent and that negative campaigning is greater the stronger his opponent's support. I differ from their analysis in allowing for the possibility of asymmetric budgets and decreasing returns to negative campaigning. More importantly, there is no empirical component to their analysis, whereas I calibrate the model to match campaign tone

data, allowing for rich heterogeneity in budget, advertising, and district-specific data.

In addition to theoretical work, there have been several structural models studying the determinants and effects of political advertising and campaigning. Gordon and Hartmann (2013a) focuses on the allocation of television advertising across markets in presidential elections. They use a BLP-type setup to understand how the Electoral College system distorts advertising decisions relative to a popular vote system. Shachar (2009) attempts to explain the finding from Nalebuff and Shachar (1999) that participation rates in U.S. presidential elections tend to be higher in states with narrower expected margins of victory. The author models campaign marketing activities in a two-candidate contest and estimates the model, finding that candidates advertise more in close states, which can drive higher turnout. Stromberg (2008) also estimates a model of the allocation of resources in U.S. presidential candidates under the Electoral College, with a focus on campaign visits rather than advertising.

Previous structural work surrounding campaign strategies differs from ours in several respects. To my knowledge, none differentiates between positive and negative campaigning, which previous empirical work has shown affect turnout and election outcomes in distinct ways.[49] Furthermore, most analyses use U.S. presidential election data and rules, in particular the Electoral College, whereas I use Senate and House elections with plurality voting systems. The focus in the campaign stage of the model is to understand the overall and relative effectiveness of positive and negative campaigning in winning elections, not on the allocation induced by electoral rules.

More broadly, the model sheds light on the overall impact of spending on elections. A few previous attempts have been made to estimate the overall impact of campaign spending on election outcomes. Palda and Palda (1998) uses regression analysis of French election data and finds a very small effect of campaign spending on vote shares. Levitt (1994) uses races with the same two candidates to estimate the effect of spending on outcomes and finds little to no effect, as well. However, Stratmann (2009) utilizes the same methodology but analyzes

---

[49]See, for instance, Ansolabehere and Iyengar (1997)

the effect of television advertising on vote shares, finding a significant impact. Additionally, da Silveira and de Mello (2011) uses a quasi-natural experiment in Brazil due to the two-round voting system and a rule that allocates TV advertising exogenously and differently in the first and second rounds. The authors find a large causal effect of TV advertising on election outcomes. Finally, in one of the few examples of structural analysis of campaign finance, Kawai and Sunada (2015) models fund-raising and spending in House elections. The authors find a relatively small effect of spending on election outcomes, slightly larger than that of Levitt (1994). For other examples, Stratmann (2005) provides a thorough review of the literature. While these works shed light on the impact of campaign contributions and spending, I focus on the effectiveness of different campaign strategies, namely negative versus positive campaigning, which can help to explain the relative difference in effectiveness for challengers and incumbents, for example.

2.3   Legal Background

In the analysis, an important component of the model is the receipts of political campaigns, taken as exogenous. Therefore, it is important to note some of the legal background surrounding campaign finance laws. While there were some changes to such laws within the sample, the regulations regarding political action committees (PAC)[50] and individual contributions did remain constant over this period. The limits on PAC contributions to candidates is the same throughout the sample. Most changes due to the Bipartisan Campaign Reform Act in 2002 involved the use of "soft money" (i.e. nonfederal funds subject to less regulation prior to the reform) and independent expenditures. Both soft money and independent expenditures deal with spending by parties and outside groups, not by the candidates. In the model, I focus only on spending by candidates, so these changes are generally not directly relevant for the type of campaign spending I consider. While outside

---

[50]Political action committees are groups "organized for the purpose of raising and spending money to elect and defeat candidates." They are an important component of campaign contributions, and face a legal limit on how much they can contribute in each cycle to each candidate.

spending may affect voters and therefore act as well, it is also important to note that within the sample, outside spending is a relatively minor part of aggregate federal campaign spending. For instance, in 2008, total spending in all federal races was approximately $5.3 billion, while outside spending comprised only about 6% of this.[51] In 2004, outside spending was only about 4.7% of total federal election spending, and in 2000 the number was 1.8%. Thus, outside spending did not play a quantitatively significant role .

In a similar vein, I also do not consider the emergence of super PACs. This is because the *Citizens United v. Federal Election Commission* Supreme Court decision, which allowed for the existence of such organizations, occurred in 2010. Super PACs can engage in an unlimited amount of spending so long as their expenditures are made independently from campaigns. Since my sample ends in 2008, super PACs do not affect the analysis.

In addition to laws regarding campaign contributions, it is also important to understand the protocol for campaign advertising. Television advertisements in the model are how I measure campaign strategies. In the United States, the main regulation on television advertisements is that stations and cable networks must "treat legally qualified candidates equally in allocating airtime." That is, if a station provides airtime to one candidate, it must offer "the same amount of airtime with the same audience size to all other candidates at the same rate," though if the other candidates cannot afford this airtime the network is under no obligation to provide it at a lower price (Karanicolas 2012). Thus, conditional on both candidates having sufficient funds, this effectively guarantees symmetry in access for each candidate in a given election. There is also a "reasonable access" rule that ensures availability of advertising time to all candidates at the rates paid by their most favored advertisers. This also implies that rationing of advertising spots will not occur in most cases. Overall, the legal rules regarding television advertising govern the setup of the model for political campaigns.

---

[51]This data is from *Opensecrets.org*, which extracts data from FEC filings. It does not accurately break down total outside spending by race type due to limitations in how the reports are filed. This is why I only list total federal election spending, which includes the presidential race in each year.

2.4 Model

I consider a model of campaigning in which there are two candidates competing for votes in an election. There are E elections held in the sample, with one Democratic and one Republican candidate in each election. Each candidate is endowed with a budget consisting of individual and PAC contributions, which is common knowledge among the candidates. Given these budgets, candidates allocate their funds between negative and positive advertising to maximize their expected vote shares. Given the chosen strategies, expected vote shares for each candidate $i \in \{1, 2\}$ are realized as a function of campaign strategies. Finally, an election-day shock that is orthogonal to information at the time campaigning decisions are made is realized, and a winner is determined.

More specifically, the model of the campaign game between candidates is similar to Skaperdas and Grofman (1995), although I deviate from it considerably. There are two candidates $i \in \{1, 2\}$ in every election. For each election, there is a population of unit mass of potential voters. Within each of these populations, a share $r_i \in [0, 1]$ initially supports candidate $i \in \{1, 2\}$, which I sometimes denote as ideological voters. These shares are restricted to be such that $r_1 + r_2 \leq 1$. The remaining share $R = 1 - r_1 - r_2$ are considered swing voters. These shares are known to candidates prior to their campaigning decisions.

Candidate i is endowed with budget B_i, comprised of all political contributions. Candidates simultaneously select the share of their budget to spend on positive or negative campaigning. I assume that different campaign strategies have differential effects on each group of voters. Negativity by candidate i primarily demobilizes his opponents' initial supporters, though they may also have the cost of demobilizing his own support, as well. These demobilized supporters do not turnout to vote. Positivity attracts swing voters. More precisely, a larger level of positivity by candidate i will attract a larger share of the R mass of swing voters, all else equal. In this sense, positivity assumed to be persuasive to swing voters.

Formally, let x_i denote candidate i's *level* of negative spending and y_i denote his positive

spending level. Given (y_1, y_2), candidate i will receive share $q^i(y_1, y_2) \equiv \frac{(1+y_i)^{1/\gamma}}{(1+y_1)^{1/\gamma}+(1+y_2)^{1/\gamma}}$ of the total mass R of swing voters, where $\gamma \geq 1$ is a parameter. These functions have the property that $q^i(y_1, y_2)$ is increasing and concave in y_i, and decreasing and convex in y_j for $j \neq i$. Furthermore, they assume that if neither side chooses any positive campaigning, swing voters will split evenly among the two candidates. Note that the smaller γ is, the more effective positive campaigning is in gaining swing voter support.

For negative campaigning levels (x_1, x_2), the total shares of support retained by candidates 1 and 2 are given by:

$$exp\{-\alpha_1 x_1 - \alpha_2 x_2\} \text{ for candidate 1}$$

$$exp\{-\alpha_1 x_2 - \alpha_2 x_1\} \text{ for candidate 2}$$

respectively, with parameters $\alpha_1 \geq 0$ and $\alpha_2 > \alpha_1$. That is, α_1 captures the idea that negative campaigning by a candidate will demobilize part of his own base, whereas α_2 reflects that negativity will demobilize part of his opponent's base.

Therefore, given campaigning choices (x_1, x_2, y_1, y_2), the expected share of support for candidate $i \in \{1, 2\}$ (with $j \neq i$) is:

$$V^i(x_i, y_i; x_j, y_j) = r_i exp\{-\alpha_1 x_i - \alpha_2 x_j\} + R\left(\frac{(1+y_i)^{1/\gamma}}{(1+y_i)^{1/\gamma} + (1+y_j)^{1/\gamma}}\right).$$

The final component determining vote shares is a mean zero random shock ε with CDF $H(\varepsilon)$. I assume ε has full support. This is an exogenous popularity shock that is unknown to the candidates when they make their campaigning decisions. It encompasses all uncertainty that is realized on election day, and is orthogonal to information the candidates have at the time they make their campaigning decisions. A given $\varepsilon > 0$ corresponds to a net gain in support for candidate 2, while $\varepsilon < 0$ corresponds to a net gain in support for candidate 1.

Therefore, candidate 1 wins if:

$$V^1(x_1, y_1; x_2, y_2) - V^2(x_2, y_2; x_1, y_1) \geq \varepsilon$$

which happens with probability:

$$H(V^1(x_1, y_1; x_2, y_2) - V^2(x_2, y_2; x_1, y_1))$$

The that probability that candidate 2 wins is thus:

$$1 - H(V^1(x_1, y_1; x_2, y_2) - V^2(x_2, y_2; x_1, y_1))$$

Given that $H(\cdot)$ is strictly increasing, the objective function is the expected vote share, $V^1(x_1, y_1; x_2, y_2) - V^2(x_2, y_2; x_1, y_1)$. Taking as given budgets, his opponent's spending decisions, and initial support, the problem of candidate 1 is:

$$\max_{x_1, y_1} \ r_1 \exp\{-\alpha_1 x_1 - \alpha_2 x_2\} + R\left(\frac{(1+y_1)^{1/\gamma}}{(1+y_1)^{1/\gamma} + (1+y_2)^{1/\gamma}}\right)$$

$$- r_2 \exp\{-\alpha_1 x_2 - \alpha_2 x_1\} - R\left(\frac{(1+y_2)^{1/\gamma}}{(1+y_1)^{1/\gamma} + (1+y_2)^{1/\gamma}}\right) \qquad (16)$$

$$\text{s.t. } x_1 + y_1 \leq B_1 \text{ and } x_1, y_1 \geq 0$$

Since the objective function is strictly increasing in both x_1 and y_1, the budget constraint binds with equality at the optimum.[52] Hence, I can rewrite the objective of the first

---

[52]I abstract from the savings/borrowing decision of the candidates across election cycles. In the sample, the median savings as a percentage of total receipts for Democrats, conditional on saving, is 1.3%, while for Republicans it is 1.4%. The median borrowing as a percentage of total receipts for Democrats, conditional on borrowing, is 2.1%, while for Republicans it is 4.4%.

candidate as:

$$\max_{x_1} \ r_1 \exp\{-\alpha_1 x_1 - \alpha_2 x_2\} - r_2 \exp\{-\alpha_1 x_2 - \alpha_2 x_1\}+$$

$$R \left(\frac{(1 + B_1 - x_1)^{1/\gamma} - (1 + B_2 - x_2)^{1/\gamma}}{(1 + B_1 - x_1)^{1/\gamma} + (1 + B_2 - x_2)^{1/\gamma}} \right) \tag{17}$$

$$\text{s.t. } x_1 \in [0, B_1]$$

The problem of candidate 2 is analogous.

A strategy for candidate i is a function x_i, which maps budgets and initial support levels to a negative campaigning proportion.[53] Formally, a strategy of candidate i therefore x_i : $\mathbb{R}_+^2 \times \triangle^2 \to [0, 1]$. That is, given budgets B_1 and B_2, and initial support levels r_1, r_2, and $R = 1 - r_1 - r_2$, $x_i(B_1, B_2, r_1, r_2, R)$ giving negativity *as a proportion of total budget* for candidate i.

The definition of equilibrium of the campaign game for a given election is as follows:

Definition 1. *Given initial support* (r_1, r_2, R), *and budgets* (B_1, B_2), *an equilibrium of this game is a pair of functions* $(\hat{x}_1(B_1, B_2, r_1, r_2, R), \hat{x}_2(B_1, B_2, r_1, r_2, R))$ *that give negative campaigning proportions for each candidate, such that for each level of initial support and budgets,* $\hat{x}_1(B_1, B_2, r_1, r_2, R)$ *solves candidate 1's problem given* $\hat{x}_2(B_1, B_2, r_1, r_2, R)$, *and vice-versa.*

2.4.1 Discussion of the Theoretical Setting

In order to make the model tractable for the empirical application, I have imposed some assumptions on the effect each type of campaigning has on voters. In this section I discuss some empirical work that supports these assumptions, as well as the potential shortcomings of the approach.

As described previously, positivity and negativity in the model differ in their effect on dif-

---

[53]Note that here I denote the strategy as a *proportion* of the total budget. This will ease notation in the calibration section, but obviously is simply a normalization since $x_i B_i$ gives the level of negative spending. Also, trivially if $B_i = 0$, then I denote x_i as 0.

ferent types of voters. Negative campaigning suppresses turnout among ideological types, while positive campaigning affects which candidate a swing voter prefers, but not her decision to turn out. One piece of anecdotal evidence on how negative campaigning can suppress turnout is from the 1998 gubernatorial elections in Minnesota, as described in the introduction. There is also empirical research that suggests negative campaigning can reduce turnout. One of the earliest studies documenting the effect of negative advertising on turnout is Ansolabehere, Iyengar, Simon, and Valentino (1994). In an experimental setting, the authors find a strong demobilization effect from negative advertising - exposure to negative ads decreased intentions to vote by 5%. They further support these findings using aggregate level data in a follow-up paper, Ansolabehere, Iyengar, and Simon (1999). Recent work by Krupnikov (2011) also supports the potential demobilizing effect of negative advertisements on supporters of the advertisement's target. Note that while there are other studies that argue that negative advertising may not demobilize as much as Ansolabehere, Iyengar, Simon, and Valentino (1994) claim (see, for instance, Finkel and Geer (1998), Ashworth and Clinton (2006), and Lau, Sigelman, and Rovner (2007)), the model does allow for demobilizing effects of negative campaigning to be arbitrarily small.

Furthermore, negative campaigning in the model not only suppresses turnout for the target candidate, but may also demobilize supporters the sponsoring candidate. This has been referred to as the "boomerang" effect in the literature. Garramone (1985) provides some of the earliest evidence for this effect, as well as the general demobilizing effect of negative campaigning. Fridkin and Kenney (2004) and Fridkin and Kenney (2011) also find that negative advertisements can depress evaluations of the target and the sponsor. However, Krupnikov (2011) argues that negative advertising may not have a demobilizing effect on supporters of the advertisement's sponsor. In calibration, I do not restrict the "boomerang" effect to be strictly positive.

2.4.2 Theoretical Results

In this section I present results from the theoretical model. Proofs of the relevant propositions and lemmas are in Appendix 2.9.1. The main result is that an equilibrium of the campaign game exists. To ease exposition, I define the following functions for $i \in \{1, 2\}$ and $i \neq j \in \{1, 2\}$:

$$MB_n^i(x_1, x_2) = r_j \alpha_2 exp\{-\alpha_1 x_j - \alpha_2 x_i\} - r_i \alpha_1 exp\{-\alpha_1 x_i - \alpha_2 x_j\} \tag{18}$$

$$MB_p^i(x_1, x_2) = \frac{2R}{\gamma} \left[\frac{(1 + B_i - x_i)^{1/\gamma - 1}(1 + B_j - x_j)^{1/\gamma}}{((1 + B_i - x_i)^{1/\gamma} + (1 + B_j - x_j)^{1/\gamma})^2} \right]. \tag{19}$$

$MB_n^i(x_1, x_2)$ is the marginal benefit of negative campaigning for candidate i computed at campaigning levels (x_1, x_2). $MB_p^i(x_1, x_2)$ is similarly defined to be the marginal benefit from positive campaigning for i. These equations are trivially obtained by taking the first-order conditions of the objective functions.

The following statements are useful in proving the existence of the equilibrium.

Proposition 1. *In the campaign stage of the model, the following statements are true:*

1. *If $MB_n^i(x_1, x_2) > 0$, then $\left.\dfrac{\partial MB_n^i(x_j)}{\partial x_i}\right|_{(x_1, x_2)} < 0.$*

2. *$\dfrac{\partial MB_p^i}{\partial x_i} > 0.$*

3. *$\left.\dfrac{\partial MB_n^i(x_1, x_2)}{\partial x_i}\right|_{(x_1, x_2)} < 0$ if and only if $\left.\dfrac{\partial MB_n^j(x_1, x_2)}{\partial x_i}\right|_{(x_1, x_2)} > 0$ and vice versa.*

4. *$\left.\dfrac{\partial MB_p^i(x_1, x_2)}{\partial x_j}\right|_{(x_1, x_2)} < 0$ if and only if $\left.\dfrac{\partial MB_p^j(x_1, x_2)}{\partial x_i}\right|_{(x_1, x_2)} > 0$ and vice versa.*

Item one of Proposition 1 state that the marginal benefit of campaigning is decreasing whenever it is positive. The first two items of Proposition 1 state that the marginal returns to both negative and positive campaigning are decreasing in the relevant range. This, as I show in the next lemma, will imply that the best response of the candidates will be singletons. Note that it is also trivial that if $MB_n^i(0, x_2) < 0$, then $MB_n^i(x_1, x_2) \leq 0$ for all $x_1 \in [0, B_1]$. If there was any x that violated this, it would be a contradiction to the fact

that MB_n^i is continuous since it would have to discretely jump from 0 to a positive value by the first item in Proposition 1.

The last two items state that when negative (or positive) campaigning is locally a strategic substitute for one candidate, it is complementary for the other. This only holds locally, and suffices to provide us with uniqueness. In the next lemma, I show that the best responses are functions.

Leamma 3. *The best response correspondence for each player, $BR_i(x_j)$, is singleton, i.e. $BR_i : [0, B_j] \to [0, B_i]$ is a function. Moreover, $BR_i(x_j)$ is continuous in x_j.*

With these two previous results, I now prove existence of an equilibrium in the campaign stage of the model.

Theorem 1. *An equilibrium of the campaign game exists.*

For the remainder of the chapter, in each election I denote the Democrat as candidate 1 and the Republican as candidate 2.

2.4.3 Empirical model

To add flexibility to the model in matching the data, I introduce heterogeneity that is unobservable to the econometrician, but observable by both candidates prior to deciding their allocations. Given parameters and budgets, the levels of initial support (r_1, r_2, R) pin down the optimal campaigning decisions. While I observe some information regarding the levels of initial support in each district or state (e.g. demographics, surveys on party support, etc.), I do not have complete information on these variables. Thus, I assume that while I can observe the *mean* levels of initial support for each election conditional on parameters, only the candidates observe the specific realization of support.[54]

More specifically, let Z denote the demographic characteristics of the district or state in which a given election is held. I assume that the initial levels of support are drawn from

---

[54]There are very few elections in the original sample that feature a prominent third-party candidate, so I do not include these in the final sample.

a Dirichlet distribution,[55] but the exact draws are observed only by the candidates. That is, the random draw of initial supports $(\tilde{r}_1, \tilde{r}_2, \tilde{R}) \sim Dir(k\overline{r}_1, k\overline{r}_2, k\overline{R})$, where $Dir(\cdot)$ is the three-parameter Dirichlet distribution, $\overline{r}_i = r_i(Z) + \psi_s S_i + \psi_{dem} D_i + \psi_{inc} Inc_i$ for $i \in \{1, 2\}$, and $\overline{R} = 1 - \overline{r}_1 - \overline{r}_2$. The functions $r_1(Z)$ and $r_2(Z)$ are known functions mapping demographics to initial support. The construction of this functions is discussed in section 2.5. k is a parameter that does not affect the mean, but is inversely related to the variance of the distribution.[56] Parameters ψ_s, ψ_{dem}, and ψ_{inc} shift the mean of the distribution. D_i is an indicator taking a value of 1 if candidate i in election e is a Democrat and 0 otherwise,[57] and Inc_i is an indicator taking a value of 1 if candidate i is an incumbent, 0 if neither candidate is an incumbent (i.e. it is an open seat election), and -1 if his opponent is an incumbent.[58] Finally, S_i is an indicator function that takes a value of 1 if candidate i is skilled and 0 otherwise. I assume that candidates observe both his own and his opponent's skill realizations, but the econometrician does not. I further assume S_i is Bernoulli distributed with:

$$Pr\{S_i = 1\} = \frac{\exp\{\beta_c\}}{1 + \exp\{\beta_c\}}.^{59} \tag{20}$$

---

[55]The Dirichlet distribution has support in the 2-dimensional simplex, making it ideal for initial support draws, since (r_1, r_2, R) necessarily must be in the 2-dimensional simplex.

[56]If $(x_1, x_2, x_3) \sim Dir(k\alpha_1, k\alpha_2, k\alpha_3)$ with $\sum_{i=1}^{3} \alpha_i = 1$, then $Var(x_i) = \frac{\alpha_i(\sum_{j \neq i} \alpha_j)}{k+1}$, while $\mathbb{E}[x_i] = \alpha_i$. Therefore, the variance of x_i is decreasing in k, but the mean is unaffected by changes in k.

[57]I allow for a shift in initial support for Democrats (i.e. ψ_{dem}) due to the fact that the measurement of Democratic support $r_1(Z)$ is systematically lower than that of Republican's. While this may be an accurate measurement of initial support, there may also be systematic undermeasurement of Democratic support. Including ψ_{dem} allows us to control for this possibility.

[58]Note that this particular structure assumes that if a candidate is an incumbent, the boost in his initial support ψ_{inc} is taken from what would be his opponent's initial support, all else equal. Results are robust to assuming that the shift in support comes from swing voters, i.e. $Inc_i = 1$ if i is an incumbent and 0 otherwise.

[59]This functional form is strictly to keep the probability of being skilled between zero and one.

Given a realization of initial support $(\widehat{r}_1, \widehat{r}_2, \widehat{R})$ and candidates' budgets, i's problem is:

$$\max_{x_i} \ \widehat{r}_i exp\{-\alpha_1 x_i - \alpha_2 x_j\} - \widehat{r}_j exp\{-\alpha_1 x_j - \alpha_2 x_i\}$$
$$+ \widehat{R}\left(\frac{(1 + B_i - x_i)^{1/\gamma} - (1 + B_j - x_j)^{1/\gamma}}{(1 + B_i - x_i)^{1/\gamma} + (1 + B_j - x_j)^{1/\gamma}}\right) \qquad (21)$$
$$\text{s.t. } x_i \in [0, B_i]$$

Since I assume the candidates observe $(\widehat{r}_1, \widehat{r}_2, \widehat{R})$, there is no change in the information set of the players, and therefore the existence and uniqueness follows. The distribution of initial support will generate a distribution of negative campaigning for each candidate, and thus will generate a likelihood function.

2.5 Data

I implement the model by using data from 2000, 2004, and 2008 House of Representatives and Senate races. In order to infer campaign strategies, I use data on political advertising tone from the Wisconsin Advertising Project. I merge this data with contribution data from the Database on Ideology, Money in Politics, and Elections (DIME), individual level opinion data from American National Election Studies (ANES), publicly available House and Senate election results, and demographic data from the 2000 Census and the American Community Survey (ACS).[60] I also hand collect some relevant data, such as incumbency status, for each race.

2.5.1 Advertising and Elections

WiscAds uses a technology that monitors the transmission of 35 national networks in the top Designated Market Areas (DMA). A DMA is a geographical region where individuals

---

[60]These data are gathered from Bonica (2013), Goldstein, Franz, and Ridout (2002), Goldstein and Rivlin (2007), Goldstein and Rivlin (2011), University of Michigan (2000), University of Michigan (2004), University of Michigan (2008), and U.S. Census Bureau (2002).

receive the same TV content and it is the smallest geographical unit in which a politician can buy air time. Every time there is a political advertisement in these markets, WiscAds captures it. A team of students research assistants then analyzes the storyboard of the advertisement to code it into the dataset. I therefore have detailed information on each advertisement: tone (i.e. whether it is positive or negative), exact date and time, station, and ad sponsor, among other things. It also importantly includes the candidate, party, or group for which the ad was aired in support. The dataset also contains an estimated cost variable. There are three ad tone types in the data: positive, contrast, and attack. I follow the convention in the literature and define negative ads as those classified as either contrast or attack ads.

The sample is limited to the geographical borders of WiscAds for each year. I merge the counties covered by WiscAds with the counties in each district.[61] Over the span of the three election cycles I consider, WiscAds should in principle cover 1,390 races. However, none of the candidates running in 814 of these elections purchased airtime and hence are not in the WiscAds dataset.[62] I therefore have no information on the campaign strategies of the candidates in those 814 elections. I also drop the 20 elections in which ads were purchased that were held in Louisiana, since this state employs a runoff system, and the 23 elections where a third-party candidate was a winner or a runner-up due to the method by which I estimate the supporters of each candidate. Finally, I drop 183 elections for which at least one candidate received positive contributions, but did not advertise, since I have no way to infer overall campaign strategies without observing advertisements. This leaves us with 361 elections over the three cycles.

Details of the type of elections covered in the final sample are given in Table 6. I have between 20 and 23 Senate elections for each year, and about 85 House elections for 2000 and 2004. For 2008, there are 126 elections included. Among the 814 elections in which neither candidate had a television advertisement, 758 were House elections. Since these

---

[61]I do not observe this directly from WiscAds. I obtained the list of counties in each DMA and year from SRDS (2000, 2004, 2008).

[62]Among these elections, 56 are Senate and 758 are House races.

Table 6

SAMPLE ELECTIONS BY RACE TYPE AND YEAR

YEAR	RACES	
	Senate	House
2000	23	82
2004	20	90
2008	20	126
Total	63	298

elections tend to be less competitive, advertising in general is less common. Hence, I observe a large fraction of the House elections where at least one candidate purchased TV ads. Within the final sample, I have 200 Republican wins versus 161 Democrat wins. The average winning margin for both parties is almost 18%. The summary of election results is available in Table 7. Table 8 shows the distribution of incumbency status in the sample. There are 16 elections for an open seat in the Senate and 61 in the House. The remaining 47 Senate races and 237 House races involve an incumbent. Given that the sample period covers a relatively successful period for Republicans, there are 180 races with a Republican incumbent, and 104 with a Democratic incumbent.

Table 7

ELECTION DATA - ALL YEARS

|	Democrats	Republicans
Winning Margin	17.75	17.82
	(14.87)	(11.61)
Winner	161	200
Total Races	361	

Table 8

INCUMBENCY STATUS

TYPE	RACES	
	Senate	House
Open Seat	16	61
Democrat Incumbent	19	85
Republican Incumbent	28	152
Total	63	298

In 2000, WiscAds followed only the top 75 DMAs, in 2004 the top 100 DMAs, and in 2008

all of the 210 DMAs, hence why the 2008 sample includes more races than previous years. Note that in each case, the DMAs cover a very large portion of the U.S. population: in 2000, the top 75 DMAs accounted for 78% of the population. In 2008, nearly the entire population is covered. However, since the observed DMAs do not exactly cover the entire U.S. population, I only partially observe the campaigns for some elections – that is, there are some races where I observe political television advertisements only in *some* of the counties within the district or state in which the election is held. To quantify the degree to which this occurs, for each race, I compute the size of the population in the intersection of the DMAs I observe and the Congressional district (for House races) or the state (for Senate races) of the election, and divide by the district or state size. I find that, on average, the dataset contains 91% of the population in a district or state. The boxplot for this measure is displayed in Figure 16. For House races, the 75th percentile is above 90% whereas for the Senate, it is around 53%. That is, for 75% of House races, at least 90% of the population is in a DMAs I observe. The median coverage for Senate races is 93.5%, while for the House it is 100%.

Figure 16
PARTY SUPPORT BOXPLOT

82
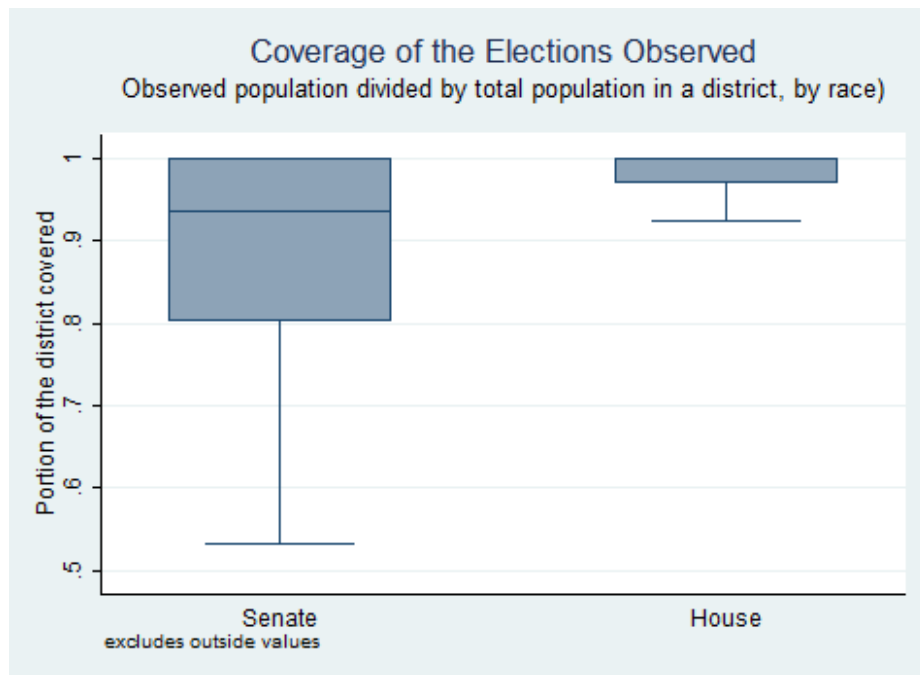
Although the dataset covers a large portion of each race, the incompleteness of the data could still be problematic for empirical implementation. The potential issue is the fact that I expect the top DMAs to contain more populous urban areas which may be more Democratic than the rest of the country. Hence, the areas I observe might have a Democratic bias, which could potentially affect the strategies of the candidates. In carrying out the empirical analysis, I assume that the candidate has the same campaign strategy across the district.

To investigate the degree to which campaign strategies may differ across different populations, I analyze the variance in campaign strategy for elections in which advertising occurs in more than one DMA. In particular, let $n_{i,e}^d$ be the cost of all negative advertisements aired in DMA d by candidate i in election e, and let $t_{i,e}^d$ be the cost of all advertisements in d aired by this candidate. I denote the campaign strategy in this DMA for candidate i in election e as:

$$N_{i,e}^d = \frac{n_{i,e}^d}{t_{i,e}^d}.$$

Letting $N_{i,e} = \frac{n_{i,e}}{t_{i,e}}$ denote the campaign strategy for candidate i in election e across *all* DMAs. Finally, I compute for each DMA the absolute deviation from the mean, $|N_{i,e}^d - N_{i,e}|$. Since air time is purchased in bulk, I consider campaigns that placed more than 500 ads in at least two different DMAs. Among these campaigns, the median absolute deviation is 0.034 for Republicans and 0.021 for Democrats. The 75th percentile is 0.093 for Republicans and 0.088 for Democrats. While there may be systematic differences between DMAs in the sample and outside the sample, this evidence is suggestive of the idea that strategies do not change dramatically across different populations.

2.5.2 Estimates of Initial Voter Support

An important determinant of the equilibrium of the model is the measure of the voter types in each election, in particular $(r_1(Z), r_2(Z), R(Z))$, where Z is the distribution of demographics in the district or state in which a given election is held in a given year.

In order to estimate those parameters, I use the joint distribution of the demographic characteristics for each district and state for the years 2000, 2004, and 2008. I construct this data using the 2000 census and the American Community Survey (ACS) for 2005 and 2008.[63] Due to data limitations I consider only race, gender, and income.

The next step is to estimate the probability an individual supports a party (or not) conditional on demographic characteristics. I use the ANES survey data to estimate the probability of identifying with a particular party conditional on demographic characteristics. In the ANES, each surveyed individual is asked about his or her relevant demographic characteristics of race, gender, and income. Furthermore, to identify party support, each individual is asked the following question:

> Generally speaking, do you think of yourself as a Republican, a Democrat,
> or an Independent? Would you call yourself a strong Democrat/Republican
> or a not very strong Democrat/Republican? Do you think of yourself as
> closer to the Republican Party or to the Democratic party?

I consider an individual to be an ideological voter if he answers this question with a strong partisan preference. The summary statistics for the ANES data are given in Tables 9, 10, and 11 for years 2000, 2004, and 2008 respectively. These data are given for the entire ANES samples, as well as broken down by party identification.

Table 9
ANES 2000 - SUMMARY STATISTICS
(BY PARTY IDENTIFICATION)

	Overall		Democrat		Swing		Republican	
	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
Black	0.119	0.324	0.236	0.425	0.077	0.267	0.023	0.149
Female	0.551	0.498	0.603	0.490	0.538	0.499	0.499	0.501
LowInc	0.518	0.500	0.575	0.495	0.513	0.500	0.443	0.497
MidInc	0.412	0.492	0.378	0.485	0.430	0.496	0.427	0.495
HighInc	0.071	0.256	0.047	0.212	0.057	0.232	0.129	0.336
# of people	1577		552		635		390	

Note: All variables dummies. *LowInc* is $[0, 50K]$, *MidInc* is $[50K, 75K]$, *HighInc* is $[75K, \infty)$

---

[63]I use the 2005 ACS for the 2004 elections since there is no 2004 ACS. Also note that, for 2008, the ACS is the three-year estimates, which allows analysis at a smaller geographic area.

Table 10
ANES 2004 - Summary Statistics
(By party identification)

	Overall		Democrat		Swing		Republican	
	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
Black	0.156	0.363	0.307	0.462	0.143	0.351	0.007	0.084
Female	0.516	0.500	0.607	0.489	0.463	0.499	0.486	0.501
$LowInc$	0.507	0.500	0.546	0.499	0.536	0.499	0.428	0.496
$MidInc$	0.404	0.491	0.396	0.490	0.385	0.487	0.438	0.497
$HighInc$	0.089	0.284	0.057	0.233	0.079	0.271	0.135	0.342
# of people	1577		552		635		390	

Note: All variables dummies. $LowInc$ is $[0, 50K]$, $MidInc$ is $[50K, 75K]$, $HighInc$ is $[75K, \infty)$

Table 11
ANES 2008 - Summary Statistics
(By party identification)

	Overall		Democrat		Swing		Republican	
	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev
Black	0.121	0.326	0.245	0.430	0.085	0.280	0.014	0.117
Female	0.545	0.498	0.620	0.486	0.488	0.500	0.535	0.499
$LowInc$	0.482	0.500	0.533	0.499	0.516	0.500	0.361	0.481
$MidInc$	0.401	0.490	0.403	0.491	0.392	0.488	0.414	0.493
$HighInc$	0.117	0.322	0.065	0.246	0.092	0.289	0.225	0.418
# of people	1577		552		635		390	

Note: All variables dummies. $LowInc$ is $[0, 50K]$, $MidInc$ is $[50K, 75K]$, $HighInc$ is $[75K, \infty)$

I estimate the probability that an individual is an ideological voter for a specific party or a swing voter using a multinomial logistic regression. I use the following variables in the estimation. $ID \in \{0, 1, 2\}$ is the party identification where 0 indicates that an individual is a swing voter, 1 indicates that the individual is an ideological Democrat and 2 a Republican. The explanatory variables for individual i in vector z_i are $(black_i, female_i, inc_i^0, inc_i^1, inc_i^2)$. The indicators $inc_0 = 1$ if the individual's income is less than \$50,000, $inc_1 = 1$ if the individual's income is between \$50,000 and \$75,000, and $inc_2 = 1$ if the income is greater than \$75,000. Setting the base outcome as being a swing voter, I estimate the vector of coefficients $\{\beta_k\}_{k=1}^2$ and get the following probabilities for each individual:

$$Pr(ID = k|z_i) = \frac{exp(\beta_k z_i)}{1 + \sum_{i=1}^{2} exp(\beta_k z_i)} \quad \text{for} \quad k \in \{1, 2\} \tag{22}$$

and $Pr(ID = 0|z_i) = 1 - Pr(ID = 1|z_i) - Pr(ID = 2|z_i)$. I separately estimate coefficients for each year. These estimation results are given in Tables 12, 13, and 14 for years 2000, 2004, and 2008 respectively.

Table 12
MULTINOMIAL LOGIT RESULTS - 2000

	Democrat		Republican	
	β_D	St. Dev	β_R	St. Dev
constant	-.51*	.261	.281	.204
black	1.3***	.178	-1.22***	.38
female	.03*	.125	-.06	.136
$0 < Inc < 50K$	-.075	.275	-.88***	.24
$50K < Inc < 75K$	-.512	.275	-.79***	.240
Observations			1577	
Psuedo R^2			.0429	
LR-χ^2			145.06	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Base outcome is Independent.

Table 13
MULTINOMIAL LOGIT RESULTS - 2004

	Democrat		Republican	
	β_D	St. Dev	β_R	St. Dev
constant	-.725	.289	.248	.234
black	1.03***	.186	-3.07***	.68
female	.674***	.157	-.20	.159
$0 < Inc < 50K$	-.125	.312	-.700**	.266
$50K < Inc < 75K$	-.064	.312	-.326	.212
Observations			1088	
Psuedo R^2			.0712	
LR-χ^2			169.26	

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Base outcome is Independent.

Table 14

MULTINOMIAL LOGIT RESULTS - 2008

	Democrat		Republican	
	β_D	St. Dev	β_R	St. Dev
constant	-.808***	.191	.376	.148
black	1.25***	.151	-1.78***	.385
female	.556***	.108	-.41***	.116
$0 < Inc < 50K$	.08	.206	-1.350***	.266
$50K < Inc < 75K$	.229	.191	-.891***	.170
Observations	2136			
Psuedo R^2	.0637			
LR-χ^2	294.75			

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Base outcome is Independent.

Finally, let Z denote the empirical joint distribution of demographics in a given election. Each element Z^i is the probability a random individual in the district has set of demographic characteristics i, where i is some combination of included characteristics. This distribution estimated from the Census and the ACS for the relevant year. Let $\overline{N} = 12$ denote the total number of possible demographic groupings. Then, I define

$$r_k(Z) = \sum_{i=1}^{\overline{N}} Pr(ID = k|i)Z^i \tag{23}$$

for $k \in \{1, 2\}$. Lastly, $R(Z) = 1 - r_1(Z) - r_2(Z)$ corresponds to swing voters in the model. Details about this variable can be found in Table 15, while Figure 17 provides the boxplot.

Table 15

IDEOLOGICAL SUPPORT FOR PARTIES - ALL YEARS

	Democratic Support		Republican Support	
	Mean	St. Dev	Mean	St. Dev
2000	0.291	0.030	0.342	0.025
2004	0.272	0.045	0.370	0.047
2008	0.259	0.038	0.385	0.037

Figure 17
ESTIMATED PARTY SUPPORT BOXPLOT
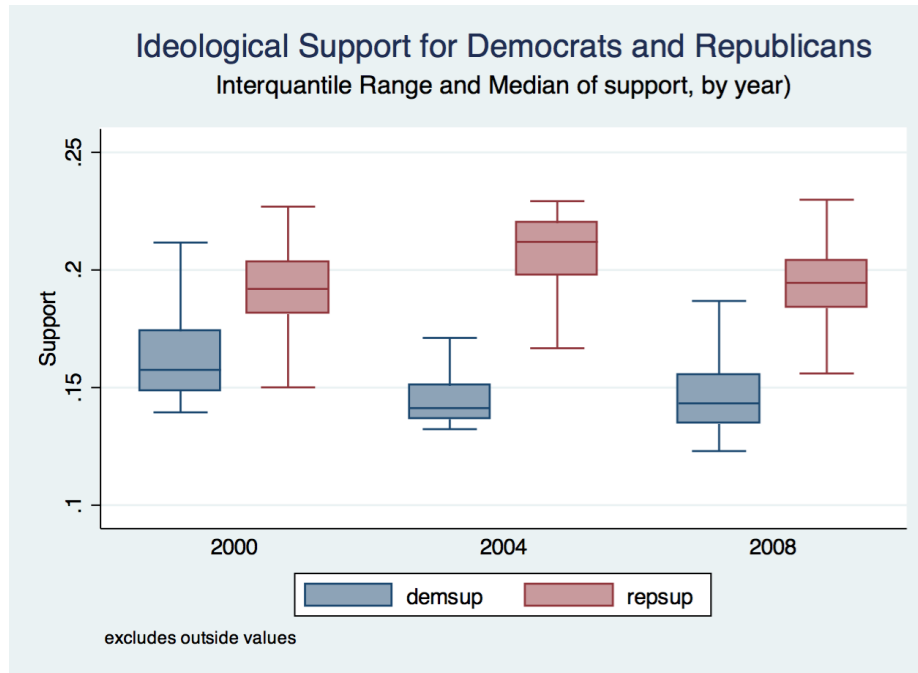
2.5.3   Candidate Budgets

I measure budgets as the total real receipts of a candidate over the campaign cycle.[64] This data comes from DIME, which extracts the receipts from Federal Election Commission filings. In the model, each candidate is endowed with a budget to allocate between positive and negative spending. While I only observe positivity and negativity for television advertising, I use this to infer overall campaign strategy. As Gordon and Hartmann (2013b) note, television advertising comprises the largest component of media spending for political campaigns. Furthermore, for both parties, television ads generally constitute a considerable element of candidate budgets, as well. Table 16 shows the total receipts by party, and the average proportion of budgets devoted to television ads in the sample is 46.0% for

---

[64]I note that, while in principle candidates can borrow or save campaign funds, in the sample saving and borrowing constitute a small fraction of total receipts. Among campaigns whose receipts exceed disbursements, the median savings rate, which is $\frac{receipts-disbursements}{receipts}$, for Republicans is 1.3% and for Democrats is 1.4%. Furthermore, among campaigns whose disbursements exceed receipts, the median savings rate for Republicans is -2.2%, and for Democrats is -4.4%. Note also that since budgets may be spent on items other than advertising, I am assuming that the tone of advertisements reflects the overall negativity of the campaign.

Democrats and 37.9% for Republicans. I therefore use receipts as the measure of candidate budgets and the breakdown of television advertising tone as the measure of campaign strategy.

Table 16
TOTAL RECEIPTS BY PARTY

	Democrats		Republicans	
	Mean	St. Dev	Mean	St. Dev
Total receipts	2,633,155	5,193,242	2,582,212	3,646,026
Ads as % of receipts	46.0%	0.403	37.9%	0.327

Note: Totals in nominal U.S. dollars.

2.5.4 Summary Statistics

I now document several of the main summary statistics and regularities in the data. Table 17 breaks down the total advertisements and advertisement tone by party. Democratic candidates placed, on average, 2,151 ads in a race, while Republicans placed about 1,963. The average number of Democratic negative ads in a race is 1,465, while for Republican candidates it is 1,274. On average, a Democratic (Republican) candidate's negative ads amount to 58.1% (54.6%) of his total ads. There is not a significant difference either in the total number of ads aired or their average negativity across parties. I observe the same pattern for the estimated costs, as seen in Table 18.

Table 17
NUMBER OF ADS AND AD TYPES BY PARTY - ALL YEARS

	Democrats		Republicans	
	Mean	St. Dev	Mean	St. Dev
Total Ads	2150.8	3440.3	1962.6	3467.3
Positive Ads	685.8	1052.7	688.9	1305.4
Contrast Ads	592.2	1092.1	443.0	925.7
Attack Ads	872.5	1979.1	830.6	1913.2
Negative Ads	1464.7	2893.3	1273.7	2536.4
% of Neg Ads	58.1%	.361	54.6%	.353

Note: "% Neg Ads" only for those with positive amount of advertising.

Table 18

Ad Costs by Party

	Democrats		Republicans	
	Mean	St. Dev	Mean	St. Dev
Total cost	1,165,178	1,920,977	1,050,046	1,745,814
Neg Ad cost	798,735	1,412,325	693,226	1,253,378
% cost of Neg Ads	58.1%	.361	54.6%	.353
Campaigns with no spending	18		13	

Note: Totals in nominal U.S. dollars.

While the broad strategies of candidates do not vary significantly across parties, the strategies do differ between incumbents and challengers, and in close races versus landslides. Table 19 provides the total ads and ad types by incumbents and challengers, whereas Table 20 does the same for estimated costs. Incumbents, on average, place about 350 more advertisements in each race and spend $250,000 more on television advertising than challengers. This stark difference is caused in part by fewer funds received by challengers. Table 21 shows that incumbents in the sample receive on average $1.1 million more than challengers. The data also show that incumbents allocate most of their air time to positive advertisements: 38.8% of incumbents' total advertising spending goes to negative ads, whereas for challengers this number is 71.5%.

Table 19

Number of Ads and Ad Types by Incumbency - All Years

	Incumbents		Challengers	
	Mean	St. Dev	Mean	St. Dev
Total Ads	2015.2	3545.9	1657.3	3353.7
Positive Ads	897.1	1433.2	348.0	682.3
Contrast Ads	368.2	840.3	536.9	1083.8
Attack Ads	749.5	2017.6	772.3	2035.6
Negative Ads	1117.7	2592.8	1309.2	2939.1
% of Neg Ads	38.6%	.352	71.5%	.306

Note: "% Neg Ads" only for those with positive amount of advertising.

Table 20

AD COSTS BY INCUMBENCY

	Incumbents		Challengers	
	Mean	St. Dev	Mean	St. Dev
Total cost	1,075,472	1,620,455	826,419	1,483,449
Neg Ad cost	609,754	1,187,704	662,548	1,280,989
% cost of Neg Ads	38.8%	0.354	71.5%	0.307
# in the data	284		284	
# with no spending	0		30	

Note: Totals in nominal U.S. dollars.

Table 21

TOTAL RECEIPTS BY INCUMBENCY

	Incumbents		Challengers	
	Mean	St. Dev	Mean	St. Dev
Total receipts	2,830,316	3,535,365	1,733,865	2,720,863
Ads as % of receipts	34.3%	0.283	41.8%	0.380

Note: Totals in nominal U.S. dollars.

Next, I classify the elections according to the ex-post vote margins and analyze the differences in advertising choices and budgets. I consider an election to be close if the winning margin is less than 5 percentage points and a blowout if the margin is larger than 20 percentage points. Then, I look at the difference between the sum of the total ads (Table 22), money spent (Table 23), and receipts (Table 24) by both campaigns. In landslide elections, of which there are 130 observations, the mean number of ads by both candidates is 1,475. In the 61 close elections I observe, the mean number of ads is 8,385, around 5.5 times as much as in landslides. Furthermore, campaigns tend to be much more negative in close elections. Around 74% of all ads aired in such elections were negative, compared to 26.5% in landslides. Similarly stark differences remain when comparing negativity in terms of money spent. Finally, as expected, total receipts in close races are much larger than in landslide elections, with around $8.1 million in the former as compared to $3.3 million in the latter.

Table 22
Number of Ads and Ad Types by Closeness of Election

	Close (<5% margin)		Landslide (>20% margin)	
	Mean	St. Dev	Mean	St. Dev
Total Ads	8385.0	9148.0	1474.5	2185.3
Positive Ads	2252.5	2726.5	891.1	1352.1
Contrast Ads	2365.3	2787.3	303.6	682.8
Attack Ads	3767.2	4606.2	279.8	663.2
Negative Ads	6132.5	6720.3	583.5	1189.2
% of Neg Ads	74.3%	0.169	26.5%	0.285
Total Elections	61		130	

Note: Totals are for both candidates in nominal U.S. dollars. "% of Neg Ads"
only for those who had positive amount of advertising.

Table 23
Ad Costs by Closeness of Election

	Close (<5% margin)		Landslide (>20% margin)	
	Mean	St. Dev	Mean	St. Dev
Total cost	4,502,892	4,694,231	708,389	1,227,679
Neg Ad cost	3,387,087	3,402,736	262,022	566,929
% cost of Neg Ads	75.5%	0.171	24.9%	0.284
# in the data	61		130	

Note: Totals are for both candidates in nominal U.S. dollars.

Table 24
Total Receipts by Closeness of Election

	Close (<5% margin)		Landslide (>20% margin)	
	Mean	St. Dev	Mean	St. Dev
Total receipts	8,114,928	10,921,429	3,331,386	5,090,576
Ads as % of receipts	65.7%	0.362	19.4%	0.190
# in the data	61		130	

Note: Totals are for both candidates in nominal U.S. dollars.

Another interesting feature of the data is the presence of corner solutions. There are many elections where a candidate's strategy is to fill his airtime solely with positive or negative advertisements. Detailed information about this, broken down by the party, is given in Table 25. In 176 of the elections, I observe both candidates allocating their air time to both positive and negative ads. For the rest, there are either no ads by one politician, or at least one candidate chooses a corner strategy. Table 26 breaks down the selected strategies by incumbency (among those elections involving an incumbent). While 97 out of the 284 incumbents in the sample chose only positive ads, 70 challengers chose exclusively negative, again reflecting the relative propensity of a challenger to campaign negatively. Only 16 incumbents went entirely negative, and only 20 challengers went entirely positive.

Table 25
DISTRIBUTION OF AD STRATEGIES BY PARTY

		Republicans				
		All positive	Interior	All Negative	No Ads	**Total**
Democrats	All positive	20	21	11	12	64
	Interior	14	176	21	1	212
	All negative	17	40	10	0	67
	No Ads	17	1	0	0	18
	Total	68	238	42	13	361

Table 26
DISTRIBUTION OF AD STRATEGIES BY INCUMBENCY

		Challengers				
		All positive	Interior	All Negative	No Ads	**Total**
Incumbents	All positive	14	29	26	28	97
	Interior	4	126	39	2	171
	All negative	2	9	5	0	16
	No Ads	0	0	0	0	0
	Total	20	164	70	30	284

2.6 Calibration and Fit

Given parameter values, for each election I can simulate draws from the initial support distribution and solve for the campaigning equilibrium. The full set of parameters is:

$$\Theta = \{\gamma, \alpha_1, \alpha_2, \beta_c, \psi_s, \psi_{dem}, \psi_{inc}, k\}$$

To select parameters, I calibrate the model to roughly match the various conditional means of negative campaigning, in particular in the aggregate and by party. I also roughly match the distributions of observed campaigning strategies. In order to calibrate the model, first consider how the various parameters differentially affect the observed distribution of outcomes. First, as the parameter ψ_{inc} increases, we will tend to observe higher initial support for incumbents and lower initial support for challengers. In turn, this will tend to generate more negativity from challengers and less negativity from incumbents. However, it will not affect negativity in open seat races. The parameter ψ_{dem} has a similar effect, except with a larger value of ψ_{dem} generating more negativity by Democrats and less by Republicans.

Next, consider the parameters affecting the marginal productivity of negative and positive campaigning, $(\alpha_1, \alpha_2, \gamma)$. Note that a proportional increase in α_1 and α_2 tends to make negative campaigning more productive. At first glance, it appears that γ could simultaneously be adjusted to keep the *relative* productivities of negativity and positivity the same, and thus not change the equilibria. Given the structure of the game, changes in γ have a differential effect on outcomes depending on the budget sizes. As an illustration, consider elections in which only one candidate has a positive budget. Without loss of generality, let candidate 1 have the positive budget. The mass of corner solutions at exclusively positive campaigning in these elections is given by the measure of $\tilde{r}_1$ and $\tilde{r}_2$ such that the marginal benefit of negative campaigning less than the marginal benefit of positive campaigning at

$x_1 = 0$:

$$Pr\left\{\tilde{r}_2 < \frac{\tilde{r}_1\alpha_1}{\alpha_2} + \frac{(1 - \tilde{r}_1 - \tilde{r}_2)}{2\gamma\alpha_2}\frac{(1 + B_1)^{1/\gamma - 1}}{(1 + (1 + B_1)^{1/\gamma})^2}\right\}.$$

Note that given the distribution of $(\tilde{r}_1, \tilde{r}_2, \tilde{R})$ is Dirichlet, which has full support on the two-dimensional simplex, this mass will be strictly positive. In the data, I observe, even among elections with only one positive budget, a wide range in values of B_1 (or B_2), as well as variation in the estimates of $r_1(Z_e)$, $r_2(Z_e)$, and $R(Z_e)$, incumbency status. For instance, for these elections the minimum budget is \$269,000 (in real 2000 dollars) while the maximum is \$3.3 million. The lowest budget is in the 6.7th percentile among all positive budgets, while the highest is in the 85th percentile. Furthermore, $r_1(Z_e)$ in elections with one budget ranges from 0.123 to 0.218, while $r_2(Z_e)$ ranges from 0.091 to 0.223. In the full sample, $r_1(Z_e)$ ranges between 0.123 and 0.282, while $r_2(Z_e)$ ranges between 0.087 and 0.230.

Since the sample of one-budget elections features wide variation in B_1 and demographics. These elections will have the same probability distribution for initial support. Consider B_1 approaching 0. The above mass of corner solutions in those elections is approximately

$$Pr\left\{\tilde{r}_2 < \frac{\tilde{r}_1\alpha_1}{\alpha_2} + \frac{1 - \tilde{r}_1 - \tilde{r}_2}{8\gamma\alpha_2}\right\}.$$

Therefore, any other combination of parameters α_1, α_2, and γ that generate the same mass should have $\alpha_1' = \kappa\alpha_1$, $\alpha_2' = \kappa\alpha_2$, and $\gamma' = \frac{1}{\kappa}\gamma$.

Now consider a similar race but with a large budget B_1'. The mass of corner solutions at $x_1 = 0$ in that election is given by

$$Pr\left\{\tilde{r}_2 < \frac{\tilde{r}_1\alpha_1}{\alpha_2} + \frac{1 - \tilde{r}_1 - \tilde{r}_2}{2\gamma\alpha_2}\frac{(1 + B_1')^{1/\gamma - 1}}{(1 + (1 + B_1')^{1/\gamma})^2}\right\}$$

Now, evaluated at the above defined α_1', α_2', and γ', I have the new mass to be

$$Pr\left\{\tilde{r}_2 < \frac{\tilde{r}_1\alpha_1}{\alpha_2} + \frac{1 - \tilde{r}_1 - \tilde{r}_2}{2\gamma\alpha_2}\frac{(1 + B_1')^{\kappa/\gamma - 1}}{(1 + (1 + B_1')^{\kappa/\gamma})^2}\right\}.$$

The last term, reflecting the marginal benefit of positive campaigning evaluated at the corner, is now more affected by the change in the γ parameter than in the low budget case, and therefore the probability mass of corner solutions will be different. Therefore, changes in the of values of $(\alpha_1, \alpha_2, \gamma)$ have a differential effect on the proportion of elections with only positive campaigning, which helps in calibrating the parameters.

Finally, consider parameters k, β_c, and ψ_s. These parameters all govern the spread in the distribution of initial support. As discussed in the empirical model section, a lower value of k corresponds to a higher variance in initial support, which generates wider variation in campaigning choices. Holding fixed incumbency status, one can think of drawing initial support from a mixture distribution, where k governs the variance of all the underlying distributions, ψ_s governs the relative means of the underlying distributions, and β_c governs the probability of drawing from each distribution.

I argue that these parameters affect the variation of initial support, and therefore campaigning strategies, in different ways. First note that, given parameters, the draw for initial support is a mixture of four Dirichlet distributions. In particular, denoting the probability of being skilled as $\tilde{\beta}_c = \frac{e^{\beta_c}}{1+e^{\beta_c}} \in [0,1]$, and ignoring ψ_{dem} and ψ_{inc} for notational simplicity:

- With probability $\tilde{\beta}_c^2$, initial support is drawn from $Dir(k(r_1(Z_e) + \psi_s), k(r_2(Z_e) + \psi_s), k(R(Z_e) - 2\psi_s))$;

- W.p. $\tilde{\beta}_c(1 - \tilde{\beta}_c)$, initial support is drawn from $Dir(kr_1(Z_e), k(r_2(Z_e) + \psi_s), k(R(Z_e) - \psi_s))$;

- W.p. $\tilde{\beta}_c(1 - \tilde{\beta}_c)$, initial support is drawn from $Dir(k(r_1(Z_e) + \psi_s), kr_2(Z_e), k(R(Z_e) - \psi_s))$;

- W.p. $(1 - \tilde{\beta}_c)^2$, initial support is drawn from $Dir(kr_1(Z_e), kr_2(Z_e), kR(Z_e))$.

Given parameters $(k, \tilde{\beta}_c, \psi_s)$, note that the mean of r_1 under the mixture distribution is

given by:

$$\tilde{\beta}_c(r_1(Z_e) + \psi_s) + (1 - \tilde{\beta}_c)r_1(Z_e) = r_1(Z_e) + \tilde{\beta}_c\psi_s, \tag{24}$$

and the variance of r_1 is given by:

$$\tilde{\beta}_c\left(\frac{(r_1(Z_e) + \psi_s)(1 - r_1(Z_e) - \psi_s)}{k+1}\right) + (1-\tilde{\beta}_c)\left(\frac{r_1(Z_e)(1 - r_1(Z_e))}{k+1}\right) + \tilde{\beta}_c(1-\tilde{\beta}_c)\psi_s^2. \tag{25}$$

To show that k, $\tilde{\beta}_c$, and ψ_s affect the distribution in different ways, I show that for $\psi_s \neq 0$, any two different sets of parameters $(k, \tilde{\beta}_c, \psi_s)$ and $(k', \tilde{\beta}_c', \psi_s')$ generate a different distribution for initial support.[65] I prove this by contradiction. Consider two different parameter values and let $\tilde{\beta}_c' = a\tilde{\beta}_c$. For the mean to be the same under both distributions, it must be that $\psi_s' = \frac{\psi_s}{a}$, as follows from (24). Additionally, let the variance under parameters $(k, \tilde{\beta}_c, \psi_s)$ as V_1 (given in (25)), and the variance under the alternative parameters be given by V_1', or:

$$a\tilde{\beta}_c\left(\frac{(r_1(Z_e) + \frac{\psi_s}{a})(1 - r_1(Z_e) - \frac{\psi_s}{a})}{k'+1}\right) + (1-a\tilde{\beta}_c)\left(\frac{r_1(Z_e)(1 - r_1(Z_e))}{k'+1}\right) + a\tilde{\beta}_c(1-a\tilde{\beta}_c)\frac{\psi_s^2}{a^2},$$

which is non-linear in both a and k. In order for $V_1 = V_1'$, I can rearrange and solve for k', which yields:

$$k' = \frac{a\tilde{\beta}_c(r_1(Z_e) + \frac{\psi_s}{a})(1 - r_1(Z_e) - \frac{\psi_s}{a}) + (1 - a\tilde{\beta}_c)r_1(Z_e)(1 - r_1(Z_e))}{V_1 - a\tilde{\beta}_c(1-a\tilde{\beta}_c)\frac{\psi_s^2}{a^2}} - 1. \tag{26}$$

Thus, it must be that if two different parameter vectors have the same distribution of initial support, with $\tilde{\beta}_c' = a\tilde{\beta}_c$, then it is a necessary condition that $\psi_s' = \frac{\psi_s}{a}$, and k' must satisfy (26).

However, the joint density of initial support under parameters $(k, \tilde{\beta}_c, \psi_s)$ can be written

---

[65]I proceed assuming $\psi_s > 0$, since if $\psi_s = 0$, then the value of $\tilde{\beta}_s$ is irrelevant as being skilled would not affect anything.

as:

$$g(r_1, r_2, R; k, \tilde{\beta}_c, \psi_s) = \sum_{i,j=0}^{1} \tilde{\beta}_c^{i+j}(1 - \tilde{\beta}_c)^{2-i-j} \left(r_1^{k(r_1(Z_e) + \psi_s i)} r_2^{k(r_2(Z_e) + \psi_s j)} R^{k(R(Z_e) - \psi_s(i+j))} \right)$$

$$\times \frac{1}{B(k(r_1(Z_e) + \psi_s i), k(r_2(Z_e) + \psi_s j), k(1 - r_1(Z_e) - r_2(Z_e) - \psi_s(i+j)))}.$$

where $B(a, b, c)$ is the beta function, and with $0 < r_1 + r_2 < 1$ and $R = 1 - r_1 - r_2$. Alternatively, under $(k', \tilde{\beta}'_c, \psi'_s)$ as specified above, the density is given by:

$$g(r_1, r_2, R; k', \tilde{\beta}'_c, \psi'_s) = \sum_{i,j=0}^{1} (a\tilde{\beta}_c)^{i+j}(1 - a\tilde{\beta}_c)^{2-i-j} \times$$

$$\left(r_1^{k'(r_1(Z_e) + \frac{\psi_s}{a} i)} r_2^{k'(r_2(Z_e) + \frac{\psi_s}{a} j)} R^{k'(R(Z_e) - \frac{\psi_s}{a}(i+j))} \right) \times$$

$$\frac{1}{B(k'(r_1(Z_e) + \frac{\psi_s}{a} i), k'(r_2(Z_e) + \frac{\psi_s}{a} j), k'(1 - r_1(Z_e) - r_2(Z_e) - \frac{\psi_s}{a}(i+j)))}.$$

Clearly, this is different from the density under the original parameters. This contradicts the supposition that the mixture distribution was identical under both sets of parameters. Therefore, changes in these parameters will affect the distributions of initial support, and therefore campaigning strategies, differentially.

Presented in Table 27 are the calibrated parameters. I note that the calibrated model implies essentially no "boomerang" effect from negative campaigning on a candidate's own supporters. I also let $\widehat{\psi}_{inc} = 0.022$, which corresponds approximately to a 4.4% advantage in initial support.[66] Additionally, the shifter $\widehat{\psi}_{inc} = 0.1157$ indicates an 11.6% increase in initial support, conditional on being skilled. The probability of being skilled, which is governed by β_c, is approximately 97%. Finally, the calibrated value of $\widehat{\psi}_{dem} = 0.045$ implies that the estimates for Democratic support based solely on demographics, $r_1(Z_e)$, are persistently low. Incidentally, the mean of $r_1(Z_e) + \widehat{\psi}_{dem}$ across all elections in the sample is 0.1987, which is approximately equal to the mean of $r_2(Z_e)$, given by 0.1958.

---

[66]This is because, by assumption, if candidate i is an incumbent and candidate j is the challenger, then mean initial support for candidate i increases by $\widehat{\psi}_{inc}$ and for candidate j decreases by $\widehat{\psi}_{inc}$.

Table 27

CALIBRATED PARAMETER

Parameter	Value
γ	1.0297
α_1	0.0001
α_2	0.5922
β_c	3.1837
ψ_s	0.1157
ψ_{inc}	0.0215
ψ_{dem}	0.0446
k	162.65

To investigate the fit of the calibrated model, Table 28 shows the average proportion of negative campaigning in the data as compared to simulations, while Figure 18 shows the distributions of negativity. The top two graphs are for Democrats and the bottom two are for Republicans, while within each pair the top presents the distribution in the data and the bottom presents the simulated distribution. Overall, the model captures several important features of the data, both quantitatively and qualitatively. In the data, overall a bit more than half (56.7%) of advertisements are negative, with Democrats performing more negative campaigning than Republicans, by about 3.8 percentage points. The model captures these features as well, only slightly predicting both overall negativity and negativity by party by approximately 4 percentage points. In addition, the data shows that challengers tend to go significantly more negative than incumbents by a margin of 71.5% to 38.8%. The model overpredicts the negativity of incumbents by only 3.7 percentage points, but more significantly underpredicts the negativity of challengers, by 8.3%. Still, the model does broadly capture the significant differences between the two groups. Finally, while the model predicts candidates in open seat elections will campaign negatively about 51.1% of the time, in the data they do so about 65.2% of the time.
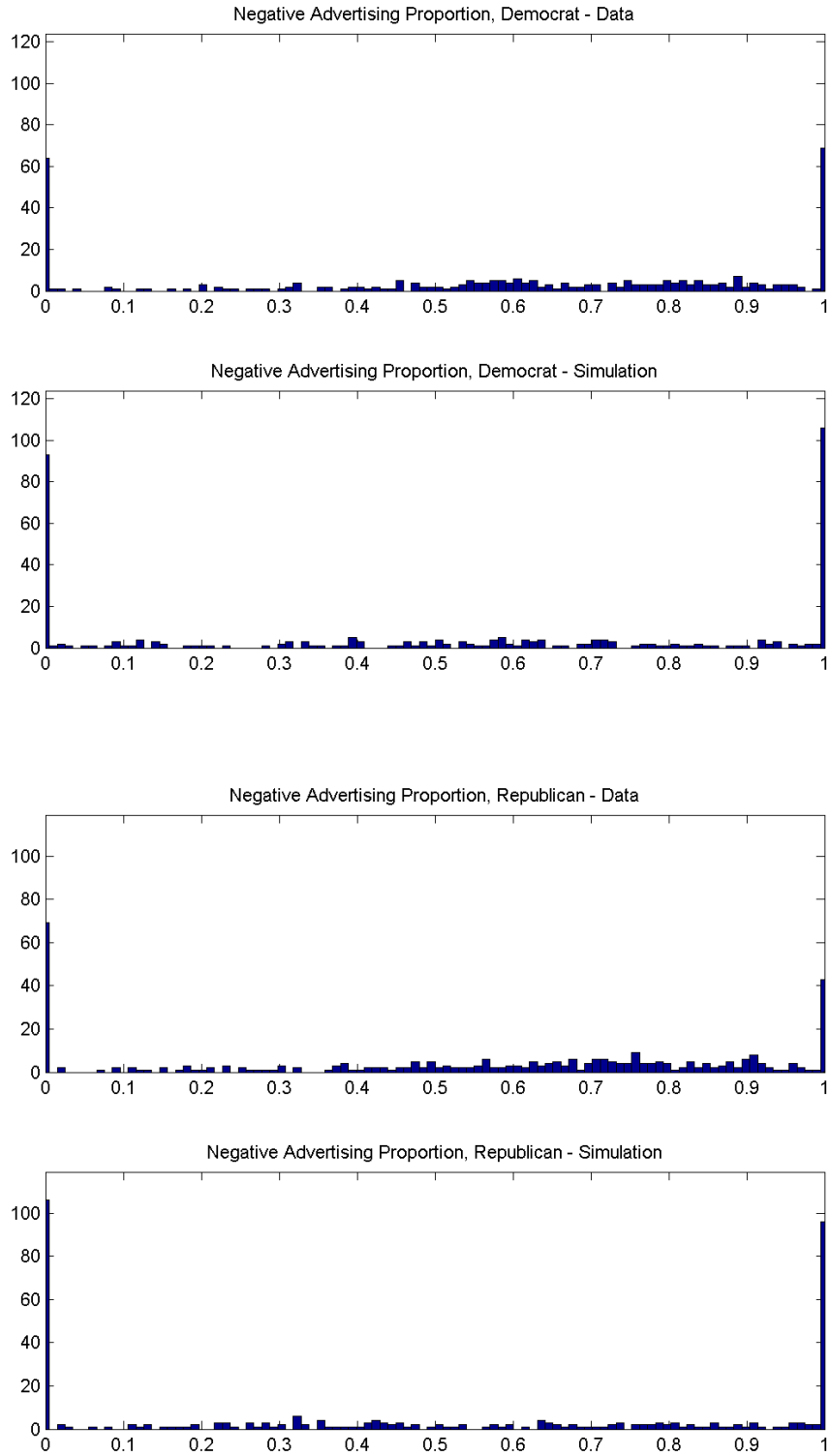
Figure 18

Histograms of Negative Campaigning Proportions, True vs. Simulated

100

Table 28

MEAN NEGATIVITY, DATA VS. SIMULATED

	Data	Simulated
Overall	0.567	0.520
Democrats	0.586	0.536
Republicans	0.548	0.505
Incumbents	0.388	0.425
Challengers	0.715	0.632
Open Seats	0.652	0.511

Note: As % of Total Budget

2.7 Results

Having shown the model can capture the salient features regarding campaign strategies, I move to analyzing the model's implications for the effectiveness of campaign strategies and spending. To interpret the remaining coefficients, I perform several exercises. As a first measure of the overall effectiveness of money, consider an open-seat election in which both candidates have the mean value of $\bar{r}_i^e$ and budgets B_i. In the sample, this implies (conditional on being skilled) values of $\bar{r}_1^e = 0.3144$, $\bar{r}_2^e = 0.3114$, $B_1 = \$2.393$ million, and $B_2 = \$2.338$ million. I then compute the change in expected vote share resulting from a 10% increase in one candidate's budgets, which corresponds to about $240,000 dollars. Note that for this exercise, I recompute the equilibrium under the new budgets. This exercise implies that, for both Democrats and Republicans, the increase in the expected vote differential in response to the increase is approximately 0.4 percentage points. This order of magnitude is consistent with Levitt (1994), which estimates that, in 1990 U.S. elections, a $100,000 increase in spending by a candidate increases his vote share by less than 0.2 percentage points for incumbents, and by between 0.19 and 0.42 percentage points for challengers. Converting $100,000 in 1990 to 2000 dollars, the model implies that such an increase in spending increases the expected vote differential by about 0.22 percentage points in an open seat elections. Unlike Levitt (1994), I find very little difference in ad effectiveness if I vary incumbency status, though at this point I hold both initial support and budgets fixed and approximately equal. Differences in marginal effectiveness across incumbents and

challengers may be largely explained by systematic differences in initial support, average budget sizes, and diminishing returns of campaign spending, which I explore below.

To analyze the model in an alternative fashion, I note that Congressional districts based on the 2000 Census contain on average 521,759 voting age individuals.[67] Under the calibrated parameter, a budget increase of $1 per voting age individual (approximately $521,759) implies an increase in expected vote difference of about 0.87 percentage points in a representative election. This is in line, albeit a bit smaller, than estimates from Palda and Palda (1998), which find using French data that "incumbent candidates can at best expect to win 1.01% of the popular vote for each extra Franc they spend per registered voter in their district."[68]

In order to more richly characterize the implications of the model, I also investigate the overall effectiveness of spending for different combinations of initial support, incumbency, and budgets. Table 29 shows results when candidate 1 is an incumbent, while Table 30 shows results for an open seat election. The tables are constructed as follows. Fixing $\overline{r}_1^e + \widehat{\psi}_{inc}$ and $\overline{r}_2^e - \widehat{\psi}_{inc}$ (which vary by column), I compute the expected vote share if there was no spending, which is given by $\overline{r}_1^e - \overline{r}_2^e + 2\widehat{\psi}_{inc}$. Then, given B_1 and B_2 (which vary by row), I compute the equilibrium of the campaign game and the resulting expected vote difference. The numbers in the tables then reflect the pre-spending expected vote difference minus the post-spending expected vote difference – that is, a negative number indicates that, after spending, candidate 2 is relatively better off. I note that B_i low is selected to be approximately the 25th percentile of all budgets, B_i mid is approximately the median budget, and B_i high is the 75th percentile, while the initial supports are analogously defined. Note also that I keep things perfectly symmetric between the two sides, except for incumbency, to control for party-specific factors.

---

[67]This estimate comes from the 2008 ACS 3 year estimates of total population over 18 by Congressional district. Note that I do not have data on voter registration data by district, so this is an upper bound.

[68]Palda and Palda (1998) used data from 1993 French elections. Converting a 1993 French Franc to 2000 U.S. dollars implies that one French Franc from 1993 is worth about $0.22 in 2000 U.S. dollars. Under the calibrated model, an increase of $0.22 per voting age individual in spending increases the expected vote difference (in the representative election) by approximately 0.2 percentage points.

Table 29
CHANGE IN EXPECTED VOTE DIFFERENTIAL FOR CANDIDATE 1, FROM NO SPENDING TO EQUILIBRIUM WITH (B_1, B_2) - 1 IS INCUMBENT

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ high	$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid	$\overline{r}_1^e$ high/$\overline{r}_2^e$ low
B_1 low, B_2 high	-2.497	-2.555	-2.761
B_1 mid, B_2 mid	0.001	-0.008	-0.131
B_1 high, B_2 low	2.528	2.497	2.452

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326, plus/minus $\widehat{\psi}_{inc} = 0.0215$ for candidates 1/2. Results in percentage points.

Table 30
CHANGE IN EXPECTED VOTE DIFFERENTIAL FOR CANDIDATE 1, FROM NO SPENDING TO EQUILIBRIUM WITH (B_1, B_2) - OPEN SEAT ELECTION

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ high	$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid	$\overline{r}_1^e$ high/$\overline{r}_2^e$ low
B_1 low, B_2 high	-2.495	-2.501	-2.619
B_1 mid, B_2 mid	0.038	0	-0.038
B_1 high, B_2 low	2.619	2.501	2.495

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326. Results in percentage points.

Table 29 illustrates several important features of the model's implications for campaign spending effectiveness.[69] First, fixing a given row, note that the percentage change decreases as $\overline{r}_1^e$ increases and $\overline{r}_2^e$ decreases. This indicates that when candidate 1 is behind, his spending is relatively more productive, consistent with previous evidence (see Levitt (1994) and Palda and Palda (1998)). This is largely driven by the fact that, when one's opponent has a high level of initial support, negative campaigning is particularly effective. This is illustrated most clearly when B_1 mid, B_2 mid and $\overline{r}_1^e$ high, $\overline{r}_2^e$ low. Here, even though both candidates have identical budgets, candidate 2 benefits relatively more from his spending, albeit a minor amount of 0.131 percentage points. More generally, table also reflects the relative ineffectiveness of spending. For instance, even when the incumbent has a high budget and the challenger has a low budget – which corresponds to a \$1.4 million advantage – in net, the incumbent's spending increases his expected margin by only 2.5%. The same dollar advantage is only slightly more effective for the challenger, yielding an increase of 2.8% for the challenger.

---

[69]Table 30 shows the same figures, except for an open seat election, with similar implications.

To investigate the relative effectiveness of positive versus negative campaigning, I consider how a large increase in either all positive or all negative campaigning affects the expected vote share. In particular, for the same combinations of budgets and initial support as above, I compute the equilibrium. Then, I compute the expected vote difference due to a sizable increase ($237,000, or 10% of the average budget in the sample) in exclusively positive or exclusively negative campaigning for one candidate, without allowing for a response from the opponent.

Table 31

CHANGE IN EXPECTED VOTE DIFFERENTIAL FOR CANDIDATE 1, FROM $237,000 INCREASE IN $B1$ - 1 IS INCUMBENT

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ high		$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid		$\overline{r}_1^e$ high/$\overline{r}_2^e$ low	
	All Pos.	All Neg.	All Pos.	All Neg.	All Pos.	All Neg.
B_1 low, B_2 high	0.4432	0.4231	0.4442	0.3744	0.4445	0.3257
B_1 mid, B_2 mid	0.4200	0.4214	0.4174	0.3744	0.4166	0.3257
B_1 high, B_2 low	0.4094	0.4109	0.3919	0.3744	0.3904	0.3257

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326, plus/minus $\widehat{\psi}_{inc} = 0.0215$ for candidates 1/2. Results in percentage points

Table 32

CHANGE IN EXPECTED VOTE DIFFERENTIAL FOR CANDIDATE 1, FROM $237,000 INCREASE IN $B1$ - OPEN SEAT ELECTION

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ high		$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid		$\overline{r}_1^e$ high/$\overline{r}_2^e$ low	
	All Pos.	All Neg.	All Pos.	All Neg.	All Pos.	All Neg.
B_1 low, B_2 high	0.4481	0.4496	0.4435	0.4042	0.4444	0.3555
B_1 mid, B_2 mid	0.4374	0.4390	0.4175	0.4042	0.4172	0.3555
B_1 high, B_2 low	0.4265	0.4281	0.3984	0.3998	0.3917	0.3555

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326.

Table 33

CHANGE IN EXPECTED VOTE DIFFERENTIAL FOR CANDIDATE 1, FROM $237,000 INCREASE IN $B1$ - 1 IS INCUMBENT, FIXED $\overline{r}_2^e$

	$\overline{r}_1^e$ low/$\overline{r}_2^e$ mid		$\overline{r}_1^e$ mid/$\overline{r}_2^e$ mid		$\overline{r}_1^e$ high/$\overline{r}_2^e$ mid	
	All Pos.	All Neg.	All Pos.	All Neg.	All Pos.	All Neg.
B_1 low, B_2 high	0.4803	0.3744	0.4442	0.3744	0.4070	0.3744
B_1 mid, B_2 mid	0.4524	0.3744	0.4174	0.3744	0.3808	0.3744
B_1 high, B_2 low	0.4247	0.3744	0.3919	0.3744	0.3667	0.3681

Note: B_i low is 0.07, B_i mid is 0.14, B_i high is 0.21. $\overline{r}_i^e$ low is 0.256, $\overline{r}_i^e$ mid is 0.291, $\overline{r}_i^e$ high is 0.326, plus/minus $\widehat{\psi}_{inc} = 0.0215$ for candidates 1/2. Results in percentage points.

Table 31 shows the results of this exercise. Note that Table 32 shows the same figures for an open seat elections, with similar implications. These results illustrate some important points. First, note that the overall vote share increases are relatively small, between 0.33 and 0.44 percentage *points*. This reinforces the notion that campaign spending is relatively ineffective at increasing vote shares. Second, note that, for fixed budget levels, the effectiveness of negative campaigning decreases noticeably as $\overline{r}_2^e$ decreases, while the effectiveness of positive campaigning remains relatively constant.[70] This is largely due to the fact that, for Table 31, I simultaneously change the initial supports for both candidates so as to keep the measure of swing voters, $\overline{R}^e$, constant. I can also increase $\overline{r}_1^e$ while keeping $\overline{r}_2^e$ fixed, which necessarily decreases the measure of swing voters. In Table 33, I show results to illustrate this. In this case, as $\overline{r}_1^e$ increases, and thus $\overline{R}$ decreases, the effectiveness of positive campaigning decreases, while for negative campaigning it remains essentially constant. This illustrates that the level of $\overline{r}_1^e$ is not the key factor for the relative ad effectiveness for candidate 1, but rather the levels of $\overline{r}_2^e$ and $\overline{R}^e$. This is particularly true since, in the calibrated model, the value of the "boomerang" effect is minor, implying that the level of own initial support is not directly important for the optimal strategy.

Finally, holding fixed a level of initial support, as B_1 increases (and B_2 decreases), I see a decline in the relative effectiveness of additional campaign spending.[71] While it is more noticeable for positive campaigning (decreasing by about 0.05 percentage points from $B_1 = \$700,000$ to $B_1 = \$2.1$ million), it is still apparent in negative campaigning. This is due to diminishing returns from campaign spending that, while not strong, are present.

2.8   Conclusion

The effect of money on election outcomes is a widely discussed topic in economics and political science. A key factor that determines the effectiveness of money and its differential

---

[70]I note that even for negative campaigning, the degree to which ad effectiveness changes as $\overline{r}_2^e$ changes is small, with changes of at most 0.1 percentage points.

[71]I also performed this exercise holding fixed B_2 as B_1 changes, and the result is essentially identical.

impact across candidates is campaign negativity, which is often overlooked by other studies. In particular, given that different candidate-types (e.g. incumbents versus challengers) use campaign funds in systematically different ways, recovering the true impact of money on election outcomes requires an understanding how effective alternative strategies are. To this end, I develop a structural model featuring a game between candidates who choose a level of negativity. Positive and negative campaigning affect different groups of voters in different ways: positivity is persuasive to swing voters deciding for whom to vote, whereas negativity affects polarized voters' decision of whether or not to turnout. Using data on levels of negativity from television advertising, candidate budgets, and other candidate- and district-specific observables, I calibrate the model, which provides implications for the overall and relative effectiveness of campaign strategies.

The calibrated model suggests that campaign spending is mostly ineffective at increasing vote shares. For the average election, which has budgets of about $2.4 million, a 10% increase in one candidate's budget increases his expected vote differential by about 0.4 percentage points. This is roughly in line with results from Levitt (1994), among others. In alternative terms, in an election where both candidates have similar levels of initial support, if one candidate has a $2.1 million budget while the other $700,000, this yields a 2.5 percentage point improvement in the expected vote differential for the first candidate. I employ other calculations to find that, albeit small, the trailing candidates benefit from extra funds more than the leading ones. I also find that negative campaigning is relatively effective for candidates who face an opponent with a high level of initial support, while positive campaigning is relatively effective for candidates in elections where neither side has a particularly high initial support. Finally, the model implies slightly decreasing returns to spending. This may, in part, explain why the previous literature tends to find challenger spending is relatively more effective than incumbent spending, as incumbents typically have large budget advantages.

2.9 Appendix

2.9.1 Proofs

Proof of Proposition 1. 1. Suppose $MB_n^i(x_1, x_2) > 0$. Then definition 18 implies that $\frac{r_j \alpha_2}{r_i \alpha_1} > exp\{(\alpha_2 - \alpha_1)(x_i - x_j)\}$. Now note that

$$\frac{\partial MB_n^i(x_j)}{\partial x_i} = r_j \alpha_2^2 exp\{-\alpha_1 x_j - \alpha_2 x_i\} - r_i \alpha_1^2 exp\{-\alpha_1 x_i - \alpha_2 x_j\}$$

which is negative if and only if $\frac{\alpha_2}{\alpha_1} \frac{r_j \alpha_2}{r_i \alpha_1} exp\{(\alpha_2 - \alpha_1)(x_i - x_j)\} > \frac{r_j \alpha_2}{r_i \alpha_1} exp\{(\alpha_2 - \alpha_1)(x_i - x_j)\} > 0$. But the first inequlity is satisfied due to the modeling assumption $\alpha_2 > \alpha_1$. The second inequality is obtained by the previous fact stated. Hence the statement is correct.

2. This is trivial since one can immediately see that $\frac{\partial MB_p^i(x_i)}{\partial x_i} > 0$ once the derivative is taken:

$$\frac{\partial MB_p^i(x_i)}{\partial x_i} = \frac{2R}{\gamma K^4} \Big[(1 - 1/\gamma)(1 + B_i - x_i)^{1/\gamma - 2}(1 + B_j - x_j)^{1/\gamma} K^2 \\ + 2K/\gamma(1 + B_i - x_i)^{2/\gamma - 2}(1 + B_j - x_j)^{1/\gamma} \Big]$$

where $K = (1 + B_i - x_i)^{1/\gamma} + (1 + B_j - x_j)^{1/\gamma}$.

3. Note that

$$\frac{\partial MB_n^i(x_j)}{\partial x_j} = r_i \alpha_1 \alpha_2 exp\{-\alpha_1 x_j - \alpha_2 x_i\} - r_j \alpha_1 \alpha_2 exp\{-\alpha_1 x_i - \alpha_2 x_j\}$$

Hence

$$
\frac{\partial MB_n^i(x_j)}{\partial x_j}
\begin{cases}
< 0 & \text{if } \dfrac{r_j}{r_i} > exp\{(\alpha_2 - \alpha_1)(x_i - x_j)\} \\[2ex]
= 0 & \text{if } \dfrac{r_j}{r_i} = exp\{(\alpha_2 - \alpha_1)(x_i - x_j)\} \\[2ex]
> 0 & \text{if } \dfrac{r_j}{r_i} < exp\{(\alpha_2 - \alpha_1)(x_i - x_j)\}
\end{cases}
\tag{27}
$$

The statement follows directly. To see it, suppose $\dfrac{\partial MB_n^i(x_j)}{\partial x_j} < 0$, that is $\dfrac{r_j}{r_i} > exp\{(\alpha_2 - \alpha_1)(x_i - x_j)\}$. Taking the inverse of both sides immediately implies $\dfrac{r_i}{r_j} < exp\{(\alpha_2 - \alpha_1)(x_j - x_i)\}$ which means $\dfrac{\partial MB_n^j(x_i)}{\partial x_i} > 0$. All other directions are similar.

4. Taking the appropriate derivatives, one can show that

$$
\begin{aligned}
sgn\left(\frac{\partial MB_p^i(x_1, x_2)}{\partial x_j}\right) &= sgn\left(-1/\gamma(1 + B_j - x_j)^{1/\gamma - 1}(1 + B_i - x_i)^{1/\gamma - 1}K^2\right. \\
&\qquad \left. +1/\gamma 2K(1 + B_j - x_j)^{2/\gamma - 1}(1 + B_i - x_i)^{1/\gamma - 1}\right) \\
&= sgn\left(B_j - x_j - (B_i - x_i)\right)
\end{aligned}
$$

where K is as defined above. The result follows immediately.

$\square$

Proof of Lemma 3. Take $\tilde{x}_j \in [0, B_j]$. First, notice that if a corner $\{0, B_i\}$ is a best response, it is the unique one. To see this, note that $0 \in BR_i(\tilde{x}_j)$ if $MB_n^i(0, \tilde{x}_j) < MB_p^i(0, \tilde{x}_j)$. But since the marginal benefit of positive ads is increasing in x_1 and that of negative ads is decreasing in x_1, $MB_n^i(x_i, \tilde{x}_j) < MB_p^i(x_i, \tilde{x}_j)$ for all x_i, which implies 0 is the unique best response. The same idea in the opposite direction applies for B_i.

On the other hand, if $x_i \in (0, B_i)$ (an interior action) is in the best response, it is the unique one. To see this note that $MB_p^i(x_1, x_2) > 0 \;\; \forall\, x_k \in [0, B_k], \quad k \in \{1, 2\}$. Since for any interior best response it must be that $MB_n^i(x_i, \tilde{x}_j) = MB_p^i(x_i, \tilde{x}_j) > 0$. Recall that when $MB_n^i(x_1, x_2) > 0$, then $\left.\dfrac{\partial MB_n^i(x_j)}{\partial x_i}\right|_{(x_1, x_2)} < 0$. Also since $\dfrac{\partial MB_p^i}{\partial x_i} > 0$, the LHS is decreasing in x_i whereas the LHS is increasing. Hence there can be only one x_i that satisfies

the condition $MB_n^i(x_i, \tilde{x}_j) = MB_p^i(x_i, \tilde{x}_j)$.

Therefore, the best response is a function and is given by

$$
BR_i(x_j) = \begin{cases} \tilde{x}_i & \text{if } MB_n^i(\tilde{x}_i, x_j) = MB_p^i(\tilde{x}_i, x_j) \\ 0 & \text{if } MB_n^i(0, x_j) \leq MB_p^i(0, x_j) \\ B_i & \text{if } MB_n^i(B_i, x_j) \geq MB_p^i(B_i, x_j) \end{cases} \tag{28}
$$

Functions $MB_k^i(x_i, x_j)$, $k \in \{p, n\}$ are continuous in both x_i and x_j. Moreover, operations $=$ and $>$ preserve continuity. Hence, BR_i must be continuous. $\square$

Proof of Theorem 1. Define the function $f : [0, B_i] \to [0, B_i]$, $f(x) = BR_1(BR_2(x))$. Obviously, a strategy profile $(x_1^*, BR_2(x_1^*))$ is an equilibrium if and only if $f(x_1^*) = x_1^*$.

By Lemma 3, both BR_1 and BR_2 are continuous, which implies that f is also continuous. Since it also maps a compact set to itself, by Brouwer's fixed point theorem, there exists $x^* \in [0, B_1]$ such that $f(x^*) = x^*$. Hence, $(x^*, BR_2(x^*))$ is an equilibrium. This completes the proof.

$\square$

Bibliography

ABRAMOWITZ, A. I. (1988): "Explaining Senate Election Outcomes," *American Political Science Review*, 82(2), 385–403.

ANSOLABEHERE, S., AND S. IYENGAR (1997): *Going Negative: How Political Advertising Divides and Shrinks the American Electorate*. New York: Free Press.

ANSOLABEHERE, S., S. IYENGAR, AND A. SIMON (1999): "Replicating Experiments Using Aggregate and Survey Data: The Case of Negative Advertising and Turnout," *American Political Science Review*, 93(4), 901–909.

ANSOLABEHERE, S., S. IYENGAR, A. SIMON, AND N. VALENTINO (1994): "Does Attack Advertising Demobilize the Electorate," *American Political Science Review*, 88(4), 829–838.

ANZIA, S. F., AND T. M. MOE (2015a): "Polarization and Policy: The Case of Public Sector Pensions," Working Paper.

——— (2015b): "Public Sector Unions and the Costs of Government," *The Journal of Politics*, 77(1), 114–128.

ASHWORTH, S., AND J. D. CLINTON (2006): "Does Advertising Exposure Affect Turnout?," *Quarterly Journal of Political Science*, 2(1), 27–41.

BAGCHI, S. (2015): "The Effects of Political Competition on the Funding and Generosity of Public-Sector Pension Plans," Working Paper.

BESHEARS, J., J. J. CHOI, D. LAIBSON, AND B. C. MADRIAN (2011): "Behavioral Economics Perspectives on Public Sector Pension Plans," *Journal of Pension Economics and Finance*, 10(2), 315–336.

BEWERUNGE, P., AND H. S. ROSEN (2013): "Wages, Pensions, and Public-Private Sector Compensation Differentials for Older Workers," NBER Working Paper.

BHATTACHARYA, S. (2012): "Campaign Rhetoric and the Hide-&-Seek Game," Working Paper.

BONICA, A. (2013): "Database on Ideology, Money in Politics, and Elections," Public version 1.0 [Computer file].

CAMERON, A. C., AND P. K. TRIVEDI (2005): *Microeconometrics: Methods and Applications*. Cambridge University Press.

DA SILVEIRA, B. S., AND J. M. DE MELLO (2011): "Campaign Advertising and Election Outcomes: Quasi-natural Experiment Evidence from Gubernatorial Elections in Brazil," *Review of Economic Studies*, 78(2), 590–612.

DIAZ-MORENO, C., AND J. E. GALDON-SANCHEZ (2000): "Collective Bargaining under Complete Information," STICERD - Theoretical Economics Paper Series 401, Suntory and Toyota International Centers for Economics and Related Disciplines, LSE.

——— (2005): "Collective Bargaining Under Perfect Information. The Negotiation of National Agreements," *Spanish Economic Review*, 7(1), 139–155.

FALCH, T., AND B. STROM (2005): "Wage Bargaining and Political Strength in the Public Sector," CESifo Working Paper Series 1629, CESifo Group Munich.

FARBER, H. S. (1978): "Bargaining Theory, Wage Outcomes, and the Occurrence of Strikes: An Econometric Analysis," *American Economic Review*, 68(3), 262–271.

FEIVESON, L. (2015): "General Revenue Sharing and Public Sector Unions," *Journal of Public Economics*, 125(3), 28–45.

FELDMAN, P., AND J. JONDROW (1984): "Congressional Elections and Local Federal Spending," *American Journal of Political Science*, 28(1), 147–164.

FINKEL, S. E., AND J. G. GEER (1998): "A Spot Check: Casting Doubt on the Demobilizing Effect of Attack Advertising," *American Journal of Political Science*, 42(2), 573–595.

FOWLER, E. F., AND T. N. RIDOUT (2013): "Negative, Angry, and Ubiquitous: Political Advertising in 2012," *The Forum*, 10(4).

FREEMAN, R. B., AND E. S. HAN (2012): "Public Sector Unionism without Collective Bargaining," Working Paper.

FRIDKIN, K., AND P. J. KENNEY (2004): "Do Negative Messages Work? The Impact of Negativity on Citizens' Evaluations of Candidates," *American Politics Research*, 32(5), 570–605.

FRIDKIN, K. L., AND P. J. KENNEY (2011): "Variability in Citizens' Reactions to Different Types of Negative Campaigns," *American Journal of Political Science*, 55(2), 307–325.

GARRAMONE, G. M. (1985): "Effects of Negative Political Advertising: The Roles of Sponsor and Rebuttal," *Journal of Broadcasting and Electronic Media*, 29(2), 147–159.

GOLDSTEIN, K., M. FRANZ, AND T. RIDOUT (2002): "Political Advertising in 2000," Combined File [dataset]. Final release.

GOLDSTEIN, K., AND J. RIVLIN (2007): "Congressional and gubernatorial advertising, 2003-2004," Combined File [dataset]. Final release.

——— (2011): "Congressional and gubernatorial advertising, 2007-2008," Combined File [dataset]. Final release.

GORDON, B. R., AND W. R. HARTMANN (2013a): "Advertising Competition in Presidential Elections," Stanford University Graduate School of Business Research Paper No. 2131.

——— (2013b): "Advertising Effects in Presidential Elections," *Marketing Science*, 32(1), 19–35.

GRIER, K. (1989): "Campaign spending and Senate elections, 1978-84," *Public Choice*, 63(3), 201–219.

GU, W., AND P. KUHN (1998): "A Theory of Holdouts in Wage Bargaining," *American Economic Review*, 88(3), 428–449.

HAO, L., AND W. LI (2013): "Misinformation," *International Economic Review*, 54(1), 253–277.

HARRINGTON, J. J., AND G. D. HESS (1996): "A Spatial Theory of Positive and Negative Campaigning," *Games and Economic Behavior*, 17(2), 209–229.

HART, O. (1989): "Bargaining and Strikes," *Quarterly Journal of Economics*, 104(1), 25–43.

HIRSCH, B., AND D. MACPHERSON (2014): "1973-2014 Union Membership and Coverage Database from the CPS," [Data file]. Retrieved from http://www.unionstats.com.

KARANICOLAS, M. (2012): "Regulation of Paid Political Advertising: A Survey," Centre for Law and Democracy.

KAWAI, K., AND T. SUNADA (2015): "Campaign Finance in U.S. House Elections," Manuscript.

KENNAN, J., AND R. WILSON (1989): "Strategic Bargaining Models and Interpretation of Strike Data," *Journal of Applied Econometrics*, 4(S), 87–130.

KRUPNIKOV, Y. (2011): "When Does Negativity Demobilize? Tracing the Conditional Effect of Negative Campaigning on Voter Turnout," *American Journal of Political Science*, 55(4), 796–812.

LAU, R. R., L. SIGELMAN, AND I. B. ROVNER (2007): "The Effects of Negative Political Campaigns: A Meta-Analytic Reassessment," *The Journal of Politics*, 69(4), 1176–1209.

Lentz, J. (2001): *Electing Jesse Ventura*. Lynne Rienner Pub.

Levitt, S. D. (1994): "Using Repeat Challengers to Estimate the Effect of Campaign Spending on Election Outcomes in the U.S. House," *Journal of Political Economy*, 102(4), 777–98.

Merlo, A. (1997): "Bargaining Over Governments in a Stochastic Environment," *Journal of Political Economy*, 105(1), 101–131.

Merlo, A., and C. A. Wilson (1995): "A Stochastic Model of Sequential Bargaining with Complete Information," *Econometrica*, 63(2), 371–399.

——— (1998): "Efficient Delays in a Stochastic Model of Bargaining," *Economic Theory*, 11(1), 39–55.

Moon, W. (2006): "The Paradox of Less Effective Incumbent Spending: Theory and Tests," *British Journal of Political Science*, 36(4), 705–721.

Munnell, Alicia H., A. J.-P. H. J., and L. Quinby (2011a): "Unions and Public Pension Benefits," *Issue in Brief SLP-20*.

Nalebuff, B., and R. Shachar (1999): "Follow the Leader: Theory and Evidence on Political Participation," *American Economic Review*, 89(3), 525–547.

Novy-Marx, R., and J. Rauh (2009): "The Liabilities and Risks of State-Sponsored Pension Plans," *Journal of Economic Perspectives*, 23(4), 191–210.

——— (2011): "Public Pension Liabilities: How Big Are They and What Are They Worth?," *Journal of Finance*, 66(4), 1207–1245.

——— (2014): "Revenue Demands of Public Employee Pension Promises," *American Economic Journal: Economic Policy*, 6(1), 193–229.

Palda, F., and K. Palda (1998): "The Impact of Campaign Expenditures on Political Competition in the French Legislative Elections of 1993," *Public Choice*, 94(1-2), 157–74.

RAGSDALE, L., AND T. E. COOK (1987): "Representatives' Actions and Challengers' Reactions: Limits to Candidate Connections in the House," *American Journal of Political Science*, 31(1), 45–81.

SANES, M., AND J. SCHMITT (2014): "Regulation of Public Sector Collective Bargaining in the States," Center for Economic and Policy Research.

SCHOLZ, J. K., A. SESHADRI, AND S. KHITATRAKUN (2006): "Are Americans Saving "Optimally" for Retirement?," *Journal of Political Economy*, 114(4), 607–643.

SHACHAR, R. (2009): "The Political Participation Puzzle and Marketing," *Journal of Marketing Research*, 46(6), 798–815.

SIEG, H., AND Y. WANG (2013): "The Impact of Unions on Municipal Elections and Urban Fiscal Policies," *Journal of Monetary Economics*, 60(5), 554–567.

SKAPERDAS, S., AND B. GROFMAN (1995): "Modeling Negative Campaigning," *American Political Science Review*, 89(1), 49–61.

SRDS (2000, 2004, 2008): *The Lifestyle Market Analyst.* Wilmette, IL : Standard Rate and Data Service.

STRATMANN, T. (2005): "Some talk: Money in politics. A (partial) review of the literature," *Public Choice*, 124(1), 135–156.

——— (2009): "How prices matter in politics: the returns to campaign advertising," *Public Choice*, 140(3), 357–377.

STROMBERG, D. (2008): "How the Electoral College Influences Campaigns and Policy: The Probability of Being Florida," *American Economic Review*, 98(3), 769–807.

UNIVERSITY OF MICHIGAN, C. F. P. S. (2000): "American National Elections Time Series Study," American National Elections Study.

——— (2004): "American National Elections Time Series Study," American National Elections Study.

——— (2008): "American National Elections Time Series Study," American National Elections Study.

U.S. Census Bureau (2002): "Census 2000," American FactFinder, U.S. Census Bureau.