The Game Theory of Government Shutdowns: Can They Be Averted?

Lee-Or Bentovim

May 28, 2025

1 Introduction and Motivation

1.1 Legislative Norms

Over the past several decades, US governmental norms have come under repeated assault by leaders from both parties. The legislative branch has continued to hand over power across a broad swath of policy concerns to the executive branch, and partisan lines have become firmer, with few prominent issues remaining cross-partisan in nature.

This is not all bad; some of the weakening of the institutions in Congress began as a necessary push to break the hold, by segregationists of both parties, on Congress, with self-evidently positive real world impacts. However, weakening institutions and norms has a long-term cost. As norm breaking becomes normalized, there are fewer and fewer sacred red lines that cannot be crossed.

For one, consider the filibuster. Although not enshrined in the Constitution, the filibuster has had a long role in the US Senate, initially requiring a two-thirds majority, later dropping to three-fifths. The filibuster as a norm worked in both ways: it was not intended as a blanket tool the minority should use to stop all legislation, but it likewise was a bulwark against strict majoritarianism. Over time, the filibuster went from an active tool that the minority party could use to hold the floor, the historic “talking filibuster”, to more of a blanket blockade used against many pieces of legislation, including the landmark Civil Rights Act of 1964, which was blocked for 60 days by a filibuster.

Over time, as the filibuster became more abused, party leaders considered the “nuclear option”, or a simple majority vote to change the rules around the filibuster: how it worked and when it could be used. In late 2013, the filibuster was removed for all non-Supreme Court nominations, under the leadership of Democratic Leader Harry Reid, and then further neutered for Supreme Court nominations under Republican Leader Mitch McConnell in 2017. In both of these cases, the norm erosion happened against a backdrop of opposition to candidates for key roles nominated by the majority party– in the short-term, each party benefited from breaking these norms and was willing to live with the unknown future costs of the other party using this new tool, in part by rationalizing that the other party would have made this change anyway.

Although there is strong evidence that norms are inherently important, the case for protecting them is even stronger when the norms protect everyday citizens from political gamesmanship among elected leaders. One prominent example revolves around government shutdowns, which occur when government funding authorization lapses. Over the past two decades, Congress has rarely passed annual budgets, instead generally opting for Continuing Resolutions (CR), especially when neither party has full control of the House, Senate, and Presidency. CRs largely keep in place the previous spending authorizations, often with minor changes thrown in benefiting one or the other party’s near-term political goals. As these funding deadlines approach, political leaders start to trade offers, often treating the proposal from the current president’s party as the starting point for negotiations. Deals are often reached, usually with only a couple of days’ worth of funding left. Sometimes, a deal is not reached, leading to permanent costs to the American economy. According to independent estimates, the 2018-19 government shutdown cost the US economy between \$3 billion at the low end and \$6 billion at the high end, in non-recoverable costs. For the individuals employed by or contracting with the federal government who are furloughed, uncertainty over when the next paycheck will arrive creates a culture of anxiety and pushes otherwise talented individuals towards working in the private sector, where they are more insulated from such arbitrary risks to their paychecks. Finally, shutdowns lead to reduced government services, such as the closure of National Parks, which impacted your author's trip in early 2019.

1.2 Self-Reinforcing Norms

Since I first interacted with Game Theory as a high school sophomore, I have been intrigued by the idea of an outside force putting its hand on the scales and changing the players' payoffs. As I entered college, furthered my interest in game theory, and witnessed the continued destruction of American norms, I began to see the destruction of norms as rational for the players, at least in the short-term. This led me to wonder if a fair, small-d democratic system could be designed such that the system punished norm breaking, calling the idea "self-reinforcing norms".

It is important to define what we mean by "self-reinforcing norms". The ideal system, according to this framework, is one where no party or actor should rationally pursue an action that involves breaking a norm. Equally, such a system should not incentivize actors to care less about norms. In the example above, which will serve as the case study of this paper, norm breaking is defined as pursuing a path which leads to a government shutdown. In order to keep this new system both fair and democratic, two further constraints must be added. First, rules added to protect norms must not be partisanly targeted, although they can be majority or minority party targeted. This means that a rule saying "if a government shutdown occurs, the Republican Party's members cannot run for reelection" is not allowed, but a rule saying "if a government shutdown occurs, members of the majority party cannot run for reelection" would be allowed. Second, rules must be social welfare maximizing. This primarily relates to scenarios where a certain outcome is not zero-sum. For example, a deal that is $(75, -75)$ is equal to a deal at $(25, -25)$ from a social welfare perspective, but if the offer of $(75, -75)$ is more likely to lead to a shutdown where utilities end up as $(-50, -50)$, then incentivizing any zero-sum deal is clearly better socially than allowing a shutdown. Only as a secondary concern might such a system aim to make the two parties equal, and then only insofar as it minimizes incentives to take norm-breaking actions.

While the initial idea that underlies this paper was intended as a three-to-four-year pursuit, for this paper we attempted to apply this approach to only one thorny issue, the aforementioned government shutdown. While our attempt at creating such a system in this paper was unsuccessful due to the structure of the game framework, I continue to believe that with further research and different frameworks, there may be pressure points which can be applied to protect institutional norms in democratic systems. When I understood that I would not be successful in finding a solution in this paper, I pivoted to simulating and better understanding the Shutdown Game which is laid out in further detail over the rest of the paper.

2 The Shutdown Game

As a general rule in US Congress, the party in power is seen as the first mover, proposing a bill which their opponents can support, oppose, or (rarely) amend. While the US House is a majoritarian institution where the majority party does not need minority party support to move bills, the US Senate is a super-majoritarian institution, where sixty votes are needed for nearly all legislation, including Continuing Resolutions to fund government in lieu of passing a full budget. Since neither party has had sixty Senators in the past 15 years, and neither party seems to be on the verge of crossing that threshold in the upcoming 2026 elections, it is fair to conclude that this will continue to remain true. While there are loopholes that allow for some legislation to avoid this path, when the majority party fails to pass a budget, CRs will continue to require sixty votes.

As the Republican Party is currently in control of all three levers involved in passing bills (House, Senate, President), we begin by plotting out a simple two-player game representing the current state of events. This game represents a simplified version of reality, but is important as a sketch of how party leaders may be viewing the current state of affairs. In the following game, the Republican Party is the column player and the Democratic Party the row player. "Unacceptable" or "Acceptable" are relative to the Democratic Party's preferences. The Republican Party, as the first mover, proposes a bill which the Democrats either accept or reject, although rejection leads to a shutdown. At this point, we consider this as a single shot game.

Republicans in Charge Game

	Unacceptable Bill	**Acceptable Bill**
Accept Bill	$(-70, 70)$	$(0, 0)$
Shutdown	$(-75, -75)$	$(-75, -75)$

Figure 1: Payoff matrix for the Shutdown Game: Democratic Party (rows) vs. Republican Party (columns)

As we can see in the below game, the Republican Party proposing an Unacceptable Bill and the Democrats accepting it is the Pure Strategy Nash Equilibrium of this game. The specific utilities at play can vary widely, but it is worth delving a little deeper into the strategies that come into play here.

Political economists often view policy outcomes as deriving from a utility function representing distance from each person's ideal point, with outcomes ranging from 0 to negative utilities, as everyone either gets what they want or has to face a trade-off. In order to make the outcomes here different, and acknowledging that debates around government funding relate generally to the status quo funding, I have strayed from that notion. An "acceptable bill" represents the continuation of the status quo. This may not in fact be (0,0), but the only assumptions we truly need to make here are **(1)** that the "acceptable bill" is worse for republicans than the "unacceptable bill", which relies solely on the assumption that from a policy perspective, what is worse for the Democratic Party is generally better for the Republican Party, which seems especially plausible when the Republicans get to propose both bills, and **(2)** that the continuation of the status quo is better for the Democratic Party than a shutdown, especially when they would certainly face blame for it when rejecting a reasonable offer.

If we accept these two assumptions, then the Republicans, when in power, can propose any "unacceptable" bill and expect that the Democrats will support it so long as it is less bad than a government shutdown would be. In the example above, the Republican proposal has left some space, perhaps due to uncertainty about the Democratic side's exact utility functions. In mid-March of 2025, Senate Minority Leader Chuck Schumer led a small cohort of Democrats in support of one such "unacceptable" bill, ensuring it passed and the government did not shut down, despite facing enormous pressure from some corners of the party to allow a shutdown. One might conclude that the bill offered to Senate Democrats was similarly close to the threshold at which a shutdown was less bad than accepting.

Before moving to the version of this game which occurs when the Democrats are in charge, it behooves us to take a moment to justify the differential negative utilities derived from government shutdowns. Broadly speaking, these fall into two categories: **(1)** do we hold the presidency and **(2)** general feelings towards shutdowns being normatively, economically, and politically bad. While the specific numbers are not very important, for the purposes of this game, we use ± 25 for holding or not holding the presidency respectively, while we set the default views towards shutdown as -100 for Democrats and -50 for Republicans, reflecting that the Democratic party, as currently constructed, places a higher value on both governmental norms and on government continuing to function as expected. Altering the payoffs changes the numbers but not the Nash Equilibrium, unless we find evidence that the true Republican cost is closer to zero than -25, at which point the equilibrium in the upcoming game (Democrats in Charge) would change[1].

Lest we believe that the strongly pro-Republican outcomes in the first game are derived solely from being the first mover, consider a second game, where the Democrats are in charge. While technically all spending bills originate from the House of Representatives, "in charge" in this context just refers to who is President, as the true determinant in modern American politics of which party holds power. In order to keep the move order consistent, note that the Democratic Party is now the column player and the Republican Party the row player. The "acceptable bill" utility is unchanged from the previous game, but the shutdown utilities have shifted based on the logic defined above, leading to a shift in "unacceptable bill" utility.

Democrats in Charge Game

	Unacceptable Bill	**Acceptable Bill**
Accept Bill	$(-20, 20)$	$(0, 0)$
Shutdown	$(-25, -125)$	$(-25, -125)$

Figure 2: Payoff matrix for the Shutdown Game: Republican Party (rows) vs. Democratic Party (columns)

[1]A case could be made that this may be true, and some future efforts may focus on pinning down this number

As in the Figure 1, the party in charge proposes a bill, and accepting a "unacceptable" bill is still the equilibrium. The immediately obvious difference between the games is that the Democratic Party gets a far smaller boost to utility here than does the Republican Party when it is in charge, due to the difference in out-party negative utility from shutdowns. This is explicitly bad per our framework: the less a party cares about a shutdown when not in office, the better they do, which incentivizes parties to not value the key norm of keeping the government open.

Importantly, this difference in outcomes is completely divorced from the magnitude of negative utility from a shutdown for the party in power– even if we believed that the party in power actually gets -1000 utility from a shutdown, the rational Democrats in the first game would still accept an "unacceptable" bill even if a shutdown harmed Republicans more than it did the Democrats, so long as it remained true that only the party in power gets to propose bills, and so long as we consider this a single shot, rather than repeated game, as we will see later.

In any case, we have at this point established that in the game as written both parties can benefit from being in power, and that shutdowns should never rationally happen in a single shot game, but that the threat of them imposed on the out-party is a factor which leads to deals being made that harm the minority party. The Democratic Party, as the party whose interests are more directly harmed by government shutdowns, is heavily disadvantaged both when in and out of power relative to the Republican Party. Less partisanly, the current structure of the US government hands more utility to the party with a lower value for the norms of keeping the government running, an outcome which in the long-term incentivizes both parties towards caring less if the government is open when they are not in power, but which also simultaneously particularly disadvantages the Democratic Party of the present moment, as the party which seeks to add legislation and increase the role of government in everyday lives has an inherently higher value on the government being in operation under regular order.

3 Influencing legislators against norm-breaking

This paper's primary attempt to protect institutions against norm-breaking relies on the fact that while parties care mostly about political outcomes and being in power, individual legislators care greatly about their own ability to be reelected. Therefore, we attempt to use the threat of ineligibility for re-election as an inducement to lawmakers to avoid norm-breaking decisions. To reiterate the requirements initially laid out in the introduction, we will aim for the following:

- Proposed rules should be impartial– they should not inherently favor one party or the other (i.e. a rule that says "Democrats are ineligible to run for re-election with $x\%$ and republicans with $2x\%$ probability" is disallowed).

- Proposed rules can punish majority or minority parties at different rates.

- A successful rule should reduce incentives to devalue norms, i.e. increase social welfare. In this specific case, a successful rule, if implemented, would change the current structure where parties benefit more the less they are hurt by shutdowns.

- A successful rule should reduce the number of shutdowns that occur.

With that out of the way, let us begin by proposing the first rule:

3.1 Uniform Ineligibility

The simplest and most straightforward version of the reelection ineligibility proposal would be a uniform penalty: If there is a government shutdown, α percent of lawmakers are ineligible to run for reelection. It is not important to define α at a specific rate, as we lack fixed numbers on the utility cost of not being able to run for office again, and even our party-level numbers on the negative impact of a shutdown are ordinal more than cardinal. We can say that in case of a shutdown, the α probability of becoming ineligible leads to an expected penalty κ, a positive constant.

To begin, let us model these changes on a variant of the party-level game we discussed earlier in section 2. Instead of considering the actors as the parties themselves, consider it as a simplified individual level model where all legislators of each party have the same utility function, thereby allowing us to incorporate the κ into the game.

Republicans in Charge Game With Uniform Penalty

	Unacceptable Bill	**Acceptable Bill**
Accept Bill	$(-70-\kappa, 70+\kappa)$	$(0, 0)$
Shutdown	$(-75-\kappa, -75-\kappa)$	$(-75-\kappa, -75-\kappa)$

Figure 3: Payoff matrix for the Shutdown Game: Democratic Party (rows) vs. Republican Party (columns)

Perhaps counterintuitively, $\forall \kappa > 0$, this rule leads to the Republican Party being able to extract even more concessions from the Democrats when in control, while the reverse is true when the Democrats are in control. Simply put, uniformly increasing the costs on the minority party has a first-order effect of increasing the space where policy concessions should rationally be made in the single iteration game, making the minority party worse off. This policy change does uniformly impact the parties, as both majority parties can extract κ more utility. If κ dominates [2], then the relative gap between the parties becomes small, but this approach does not lead to more reasonable and responsible policy, just a massive swing based on who is in control in our present framework.

Before delving into more complex proposals, it is important to note that beyond the obvious assumptions that went into aggregating all lawmakers of a party into having the same utility, one other key assumptions is that all lawmakers are planning on running for reelection,

[2] In my public policy class, I am analyzing this problem from the approach that takes κ as a dominant force, creating a kind of brinkmanship problem. I don't think the math is that interesting, but the idea is that if the cost of playing chicken goes up, the expected utility from playing chicken is negative and so no one will ever choose to play. I am unsure which result I believe, but this is an interesting case of changed framework leading to different conclusions

which is true for most, but certainly not all lawmakers[3]. Still, it seems reasonable to assume that a roughly equivalent proportion of lawmakers from each party are not planning on running for reelection, allowing the general framework to hold up.

3.2 Majority v. Minority Differential Ineligibility

Moving up one level of complexity, our framework explicitly allows for differential penalties on the majority and minority parties. Using the same κ from earlier, we can represent this as κ_{maj} and κ_{min}. We can now show a new variant of our existing party level game.

Republicans in Charge Game With Majority and Minority Penalty

	Unacceptable Bill	**Acceptable Bill**
Accept Bill	$(-70-\kappa_{min}, 70+\kappa_{min})$	$(0,0)$
Shutdown	$(-75-\kappa_{min}, -75-\kappa_{maj})$	$(-75-\kappa_{min}, -75-\kappa_{maj})$

Figure 4: Payoff matrix for the Shutdown Game: Democratic Party (rows) vs. Republican Party (columns)

One key difference that emerges here is that the majority party can only move policy in their direction by a further κ_{min}, not κ_{maj}, as the determining factor remains the preference of the minority party. However, this proposal still leaves the minority party in a worse (or if $\kappa_{min} = 0$ the same) position as they were in prior to implementing this rule. Additionally, given that we cannot change κ_{min} based on which party is in power, this proposal also does not help us punish parties that inherently care less about shutdowns occurring.

The emerging through-line between these two games, which remains true for all variants of rules that modify only the shutdown utility, is that $\forall \kappa > 0$, adding a cost to shutdown hurts the minority party, but also hurts it by the same κ or κ_{min} factor regardless of who is in charge, thereby doing nothing to decrease the gap between party utility when in charge of first-moving in the game. Therefore, if our goal is to stop incentivizing parties to care less about government shutdowns, pushing further down this electoral penalty seems unlikely to be fruitful.

Instead, we can consider that while each shutdown game is a discrete act, it occurs between two players with memories and the knowledge that there will be future such games; the shutdown game operates as a repeated game. Analyzing it as such can help expose repeated game strategies that work well and not so well, better preparing us for a further attempt at norm-reinforcement.

4 Repeated Game Simulation

In order to better evaluate the shutdown game, we created a simple model representing the players, the game, and various strategies players might employ. GitHub code for this can be

[3]Really this subset of legislators who are not planning on running for reelection have significantly altered utility functions regardless of κ, which could be interesting to explore in its own right

found here. In the following sections, we will breakdown the structure of the game, several strategies available to the players, and discuss some simulated results, and what conclusions we can draw about this game based on them.

4.1 Game Structure Definition

The game class defined in the code is largely the same as the simple model defined in section 2, with the largest change being that rather than "acceptable" and "unacceptable" offers, the proposer can make offers over a predefined range, defaulting to $[0, 100]$. In order to make the game repeated, we've added a probability p, which represents the probability of proposer and responder flipping, as if a new majority party was elected. In simulations we set $p = 0.25$, a reasonable recent estimate of the odds of the Presidency flipping in a given election. The simple strategies defined below do not think about p, making it irrelevant in the long-term, although changing p does impact risk-aversion in strategies that are aware of it.

The game takes as input two players, each of whom has both a proposing and responding strategy which they use when in each role. In simulations we defaulted to having the Republican player lead at first, but this is irrelevant in the long run with a sufficiently large p. Players in the game were not given access to the other player's utility functions or strategies, although some strategies allowed attempts to learn the opponent's utility indifference point.

4.2 Proposer Strategies

We define several proposer strategies, each of which represents a simplified version of an approach a player in this game may take over a repeated game. Each strategy considers and cares about different inputs and outputs to determine best proposals.

1. **Conceding Proposer**: The Conceding Proposer starts at the highest offer available and decrements its offers until it is accepted. If its proposal is accepted, it will continue to offer the same proposal. This strategy approximates a simple effort to find, over the long term, the best offer for the proposer that will always be acceptable to the responder. Against several responder strategies this strategy leads to an equilibrium that looks remarkably similar to the single shot game equilibrium. Against other strategies it performs remarkably poorly as it is very trusting that the opponent is being short-term rational.

2. **Random Proposer**: The Random Proposer has two variants, it can either be uniformly random or normally distributed around an inputted mean with an inputted standard deviation. It selects a proposal from within the given distribution without any regard for what the opponent is doing. While this may seem sub-optimal, it is strategy proof: the responder has no opportunity to send costly signals to the proposer, and if the responder is able to change strategies when it recognizes this, random proposer should be able to get them to switch to utilitarian responses[4]. In simulations

[4] I unfortunately did not have the opportunity to build in switching strategies mid-game, although this claim is accurate and a main draw for a random proposer

without strategy changing, Random Proposer was a loser since responders were stuck with their strategic choices.

3. **Tit-for-Tat Proposer**: The tit-for-tat proposer starts out by offering the "acceptable" offer of $(0, 0)$, and will continue to do so until the opponent deviates, at which point it will always copy the opponent's most recent offer. When both players play tit-for-tat, we arrive at the socially optimal strategy, given that all responder strategies will accept the offer of zero, the best possible offer.

4. **Learning Proposer**: This was the first learning strategy I developed, but it was very jumpy, not settling at the right location, and was (mostly) dominated by Binary Search Proposer. The idea here was to decrease offers if rejected and increase offers if accepted, to continually try and gain the most advantage possible. If player utility functions were less fixed than they are in the construction of this game, this strategy may well outperform Binary Search Proposer.

5. **Binary Search Proposer**: This learning strategy tracks the range in which an accepted offer may fall, attempting to prevent itself from offering a deal the opponent will reject, while maximizing the utility it can get from its opponents. This is an improved version of Conceding Proposer, and therefore also arrives at a stable equilibrium resembling the one shot game equilibrium against a utilitarian responder, although it could use some further reworking against other strategies, where there is a potential for min and max to become inverted when the opponent plays non-utilitarianly.

6. **Risk Aware Proposer**: This proposer strategy is the first that is aware of the downsides of certain proposals, unlike its predecessors which are only trying to maximize upside. It makes the offer with the highest expected utility, hard-coded for now into intervals of ten between 0 and 100, and stores outcome data to continue to hone in on the best guess. It does not consider the future value from certain proposals, just the present value.

7. **Forward Looking Proposer**: This proposer was the only one designed to consider the future costs associated with making a certain proposal. It was designed to test the idea of p, the probability of control of the game flipping, being increased when a shutdown happens. Unlike Risk Aware Proposer, Forward Looking Proposer considers not just the probability that an offer is rejected, but also that rejection may lead to an increased risk of losing control of the game. While the coded version is relatively simple and falls short of truly responding to this, this framework is the most interesting to consider moving forwards (see section 4.5).

While overall analysis will be reserved until after discussing the responder strategies, given the unchanging utility functions, most of these strategies landed in similar locations against a wide swath of responders, generally stabilizing near or slightly below the maximum utility they could extract from the responders, with exceptions being the two strategies that did not consider the opposing responders, tit-for-tat and random.

4.3 Responder Strategies

On the responder side, we likewise define several strategies that players could choose to utilize. As with the proposers, different responder strategies performed better against different proposers, and tracked different pieces of information.

1. **Utilitarian Responder**: This simple responder strategy merely compared the utility of accepting and rejecting the offer and accepted the one that gave it better utility. This strategy is the "rational" strategy for a one shot game, but did not prove to be the best response for the most part against strategies in the repeated game.

2. **Tit for Tat Responder**: This responder accepts any deal that is as good as or better than the previous offer made to it by the other player. It therefore incentivizes stabilizing and punishes strategies that try and find the maximum it is willing to accept, like Binary Search responder. In order to ensure it doesn't accept the first offer it sees no matter what, it only considers acceptance if accepting would be the utilitarian choice.

3. **Probabilistic Responder**: This responder takes a probabilistic approach to accepting offers. It takes an input alpha which impacts how sensitive it is to utility difference between accept and reject thresholds, with values closer to zero lead to higher variance in behavior. It uses a logistic sigmoid to calculate the likelihood of accepting the offer. This responder outperforms Utilitarian responder against most opponents, getting most proposers to stabilize at a meaningful level below the utilitarian break-even point. We go into further detail in Section 4.5, but this strategy is the most realistic strategy a party in this game might consider, given that they need to balance both short- and long-term interests.

4. **Strategic Rejector Responder**: This responder strategically rejects offers to motivate the proposer to offer better deals. This implementation is quite naive, it accepts all offers that are equal to or better than the previous offer, so long as accepting is utilitarian. If it sees the same offer again, on the fifth consecutive occurrence it will reject an equal offer as well. If there is a known floor, over time Strategic Rejector will move some proposers to making ideal offers, if the floor is unknown but exists, Strategic Rejector will reject offers forever. As a note, a proposer which occasionally makes worse offers would be able to reset Strategic Rejector by making a much higher offer and then returning to a lower one that is still higher than the preceding offer. As mentioned in the Probabilistic Responder section, a Strategic Rejector in this exact structure would probably face too much short-term gain to live to see the light of the long-term benefits, although it is useful to highlight that willingness to take losses certainly influences future behavior from a cross-section of proposers.

In most cases, Probabilistic Responder outperformed the others with the exception of the Strategic Rejector, which felt almost like cheating based on how it was defined. By committing to this strategy, players were often able to nudge their opponents to better and better deals. This illustrates a common trope of repeated play games, that taking a strategic loss is often a wise strategy to help in the long-term.

4.4 Strategy Tournament

After creating all of these proposers and responders, we ran a complete round robin tournament randomizing for both players which proposer and responder strategy they were assigned. With seven proposers and five responders (we used both a normal and a high variance version of Probabilistic Responder), there were $35^2 = 1225$ pairings, each of which was run for 50 different matches each lasting 200 rounds. We ran this tournament under the basic rules defined throughout this paper.

4.4.1 Republican Highlights

For the Republicans, the best performing approach was Forward Looking Proposer paired with Strategic Rejection Responder, making up the top six performing performances. Against a variety of Democratic proposers and responders, this led to as high as 6000 utility over 200 trials, averaging over 30 per round, despite spending about half of rounds in the minority where they could do no better than zero. This pair of strategies was notably left-skewed, with a mean performance of 2000 and a median performance of 4000 against all Democratic strategies. Other highly performing Republican strategies included Risk Aware Proposer and the high variance version of Probabilistic Responder. Unsurprisingly, the top performing strategies generally had 20-30 rejections over a 200 turn game, compared with a mean and median closer to 70.

A graph of utility over time for the highest Republican outcome can be seen here:

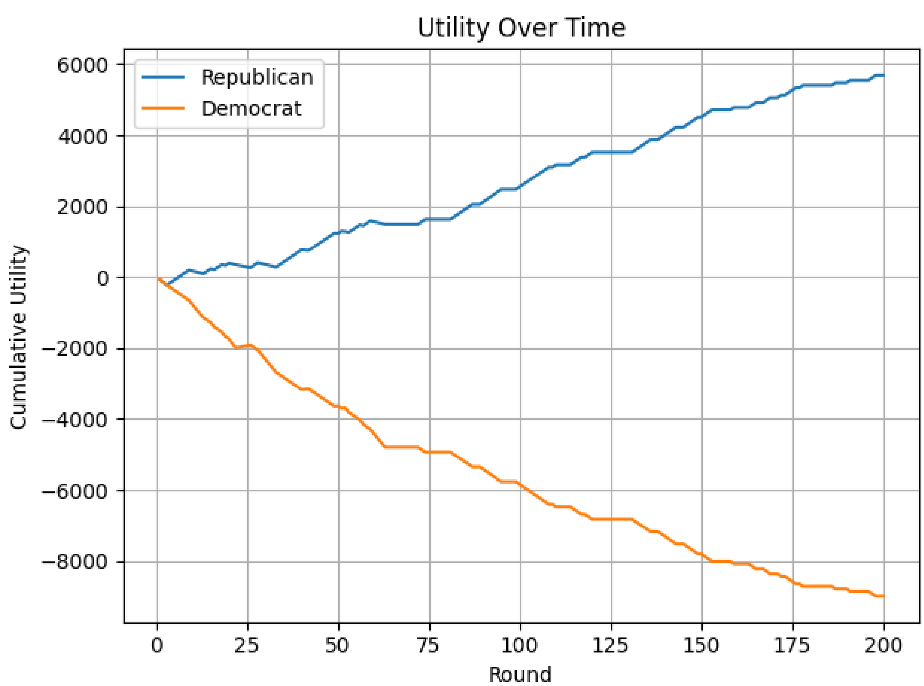

Figure 5: Example Run: Best Republican Outcome

On the other end, some strategies did not go so well for the Republican player. Of the bottom fifty results, all of them involved the Democrat being a Strategic Rejector Responder against the Republican's Binary Search Proposer or Tit-for-Tat Proposer. Notably, the other half of the equation, i.e. the Democrat's proposal strategy and Republican's responder strategy, were quite varied across these worst outcomes. Less surprisingly, the number of rejections was incredibly high, reaching as high as $\frac{192}{200}$ in the worst case scenario, but remaining above 90 rejections in all of the bottom fifty. In several of these scenarios the Republican player even performed worse than the Democrat player despite the structural advantages given to them. Republican utility in these games was between -7000 and -9000.

A graph of utility over time for the worst Republican outcome can be seen here:

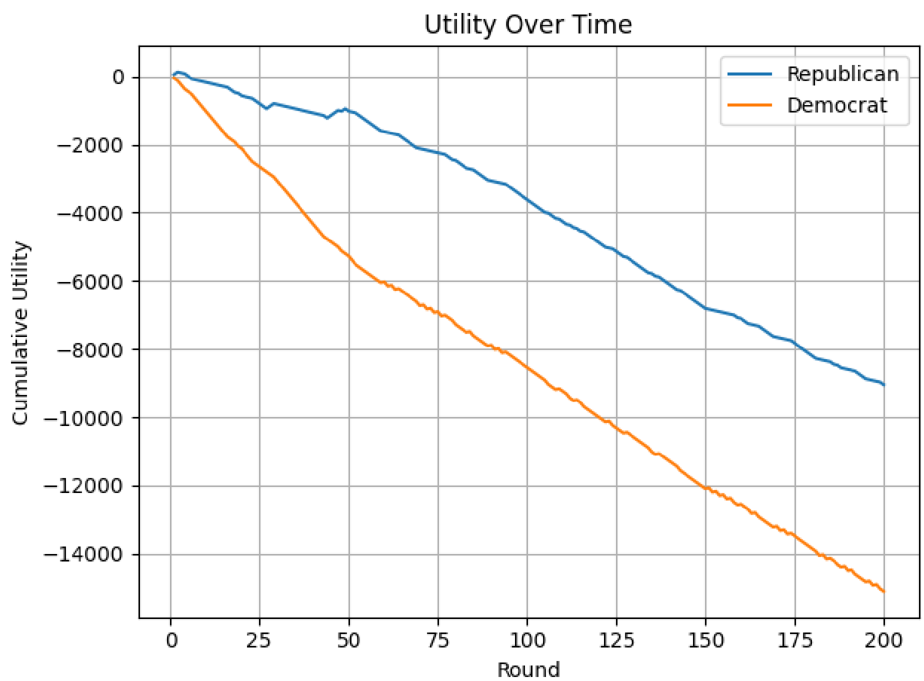

Figure 6: Example Run: Worst Republican Outcome

4.4.2 Democratic Highlights

On the Democratic side, success was far more limited due to the constraints of the game. The best performing Democrats occurred when both players were Tit-for-Tat Proposers, since both would offer zero at all times, which was always accepted. When we exclude Tit-for-Tat (either both or even just whenever the Republicans played Tit-for-Tat), the best performing Democrats came when they played Strategic Rejector Responder against a Conceding Proposer, and thus forced the Republicans down to offering zero over time. Success also came occasionally against Republicans making Random proposals. Unsurprisingly, successful Democrats saw very few rejections, even lower than in the cases with the best Republican performance, remaining in the $10-20$ rejection range for the most part. Finally, even the "best" Democrat, when excluding Tit-for-Tar scenarios received -2700 utility over the course of the game, showing that this game was tilted strongly against the Democrats, as the party with higher rejection costs.

A graph of utility over time for the best (non Tit-for-Tat) Democratic outcome can be seen here:

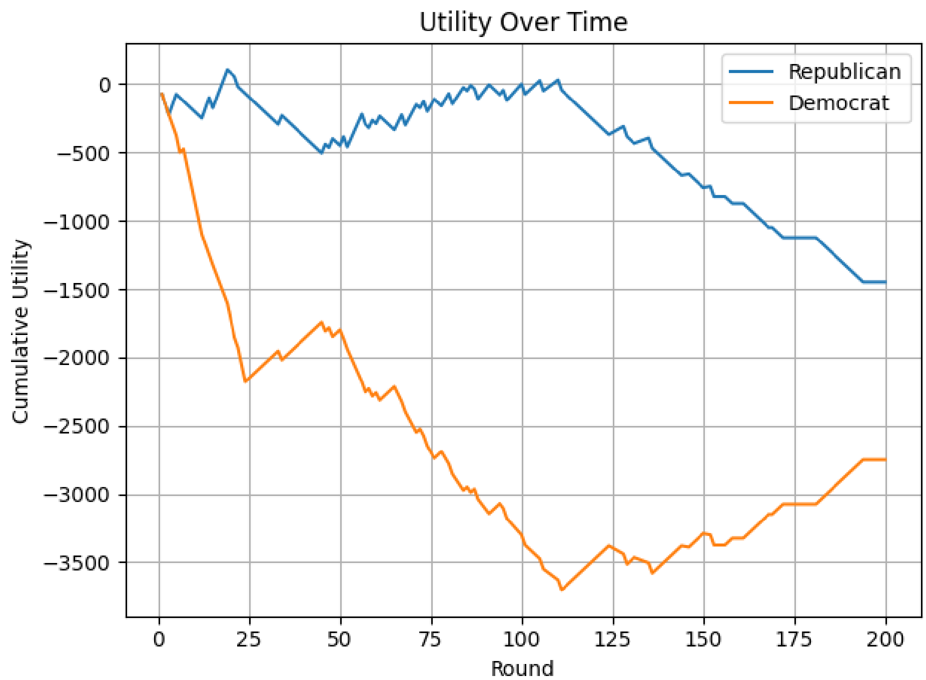

Figure 7: Example Run: Best Democratic Outcome

On the other extreme, this game could go incredibly poorly for the Democrats with a bad combination of strategies. Tit-for-Tat proposer was also prominent at this extreme, with unlucky Democrats using this strategy against well calibrated Republicans performing incredibly poorly. In the worst case, these Democrats got over $-20,000$ utility, averaging a loss of more than 100 per round, more than a Republican could theoretically lose. Having the wrong proposer strategy seemed to be the biggest risk to the Democrats, as after Tit-for-Tat, dozens of bad Random Proposer matches followed, trailed by several more Binary Search Proposers and Learning Proposers. For the most part what these had in common was a very high rejection rate, over 50%. Unlike the best performing Democrats, which often came from circumstances out of their control, the worst performing Democrats pretty consistently used strategies that proved to be very poorly suited to the game from their perspective.

A graph of utility over time for the worst Democratic outcome can be seen here:

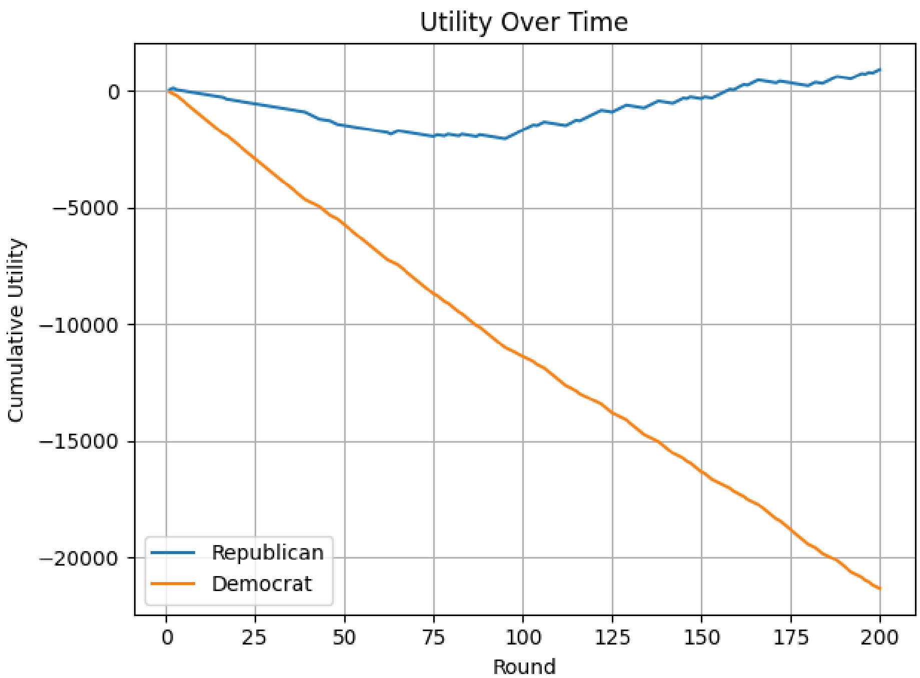

Figure 8: Example Run: Worst Democratic Outcome

4.4.3 Relative Performances

Next, it is interesting to consider some relative performances: when did the Republicans do best relative to the Democrats and vice versa. While utility is generally not relativistic, exploring which strategy quadruples led to extreme outcomes can be illuminating. Beginning with Democratic over-performance, the Democrats beat the Republican in 27 of the 1225 matchups, less than 2% of the time. All of these Democrats were Strategic Rejector Responders, although the other three roles were not fixed. On average there were around 90 rejections in games of this type, suggesting that the approach here was Democrats being successfully at “I’ll drag you down with me”. Utilities in these games were in the −5000 to −8000 range, better for Democrats than their average. While we did not build the ability for this, Republicans in such a situation likely would have benefited from changing their approach.

A graph of utility over time for the best relative Democratic outcome can be seen here:

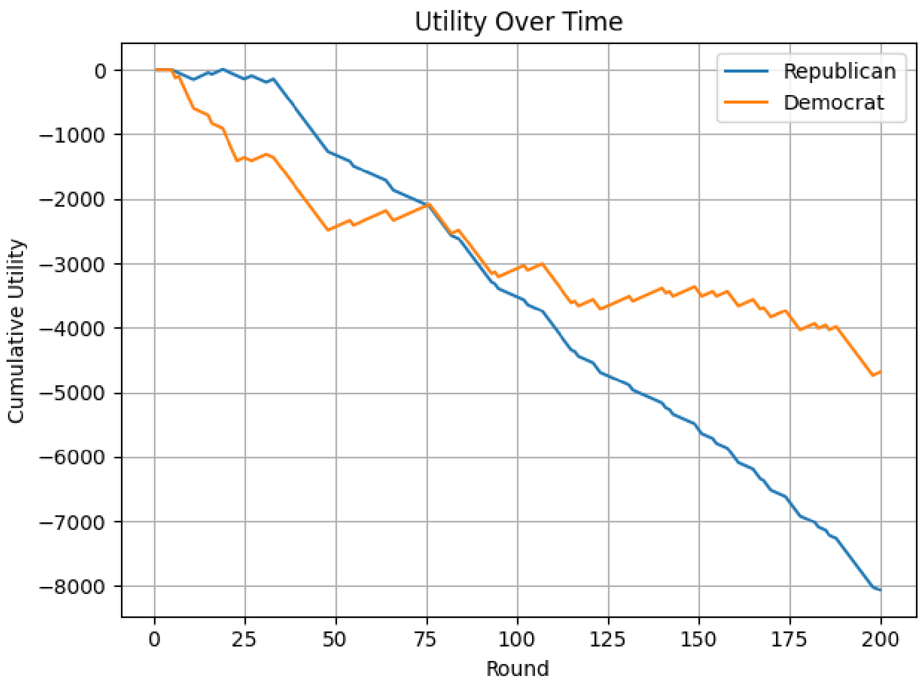

Figure 9: Example Run: Best Democratic Relative Performance

On the other end, the best relative Republican performances came from Democrats playing Tit-for-Tat Proposer, although again the other three roles varied. Many of the Democrats in this case were making offers that the Republicans were always going to reject while accepting pretty bad deals when the tables were turned. Thus, the Republicans in these cases performed a third to a half worse than in their best case outcomes, staying in the $3000-4000$ range, while the Democrats were close to their floor, mostly in the $-19,000$ range. Roughly half of offers were rejected in these games, mainly represented as rejection of all of the Democratic proposals and vanishingly few of the Republican ones.

A graph of utility over time for the best relative Republican outcome can be seen here:

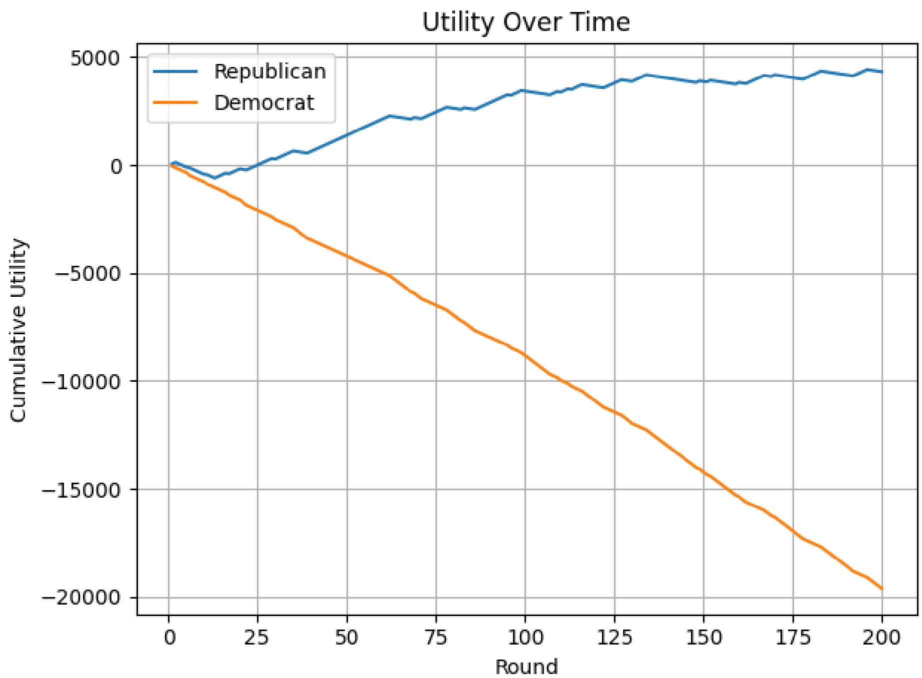

Figure 10: Example Run: Best Republican Relative Performance

4.4.4 Simulation Conclusions

Unsurprisingly given the structure of the game, the Democrats never received positive utility after 200 iterations. Their best case scenario relied on both parties playing the altruistic (social welfare maximizing) approach with Tit-for-Tat proposers, but while this was great for the Democrats it was towards the lower end of Republican outcome distribution. Democrats choosing this approach also took a great risk: against other Republican strategies Tit-for-Tat was the worst strategy available. The Republicans, given their structural advantages in this game, had more control over their best outcomes, but there was no dominant strategy for the Republicans either. More broadly, there was not an equilibrium to the game: in a real world, both players would want to update their strategies as they better understood their opponents.

4.5 Going Beyond the simulation

Moving beyond the results of the simulation, the path towards a realistic equilibrium explanation likely relies on closer consideration of a game between the two parties where both are Forward Looking Proposers and Probabilistic Responders. While the coded version of the Forward Looking Responder is not quite where I would want it to be, I think this pairing best represents real world strategies that the players might employ. For the proposal strategy, leaders are thinking not just about the current game, but about the next several iterations likely to arise. In a scenario where the probability of being the next proposer is impacted by rejections, this encourages caution from the majority party. On the other end, the real world responder party will struggle to act as a true strategic rejector, given the short-term losses will mount too much for the party to bear. At the same time, being strategically unpredictable can have benefits. It is harder for the opposing party to maximize the utility they can extract from you if they cannot determine where your boundary lies.

In this subsection we will go into further detail of the defined strategies and explore a model pitting these strategies together[5]. We will use p as the probability of control flipping, p_reject_bump as the boost to p when the previous offer was rejected.

While the Forward Looking Proposal can get very complicated very quickly, a simplified version wants to select x to maximize expected utility over n terms. Attempting to write this out for just the first two turns shows how quickly this can get out of hand, but also provides the idea for what can be maximized.

$$EU_x^0 = x_0 \cdot P(\text{Accept}(x_0)) + u_{bad} \cdot (1 - P(\text{Accept}(x_0))$$

The expected utility in the current term is the probability of the opponent accepting the offer multiplied by the utility you would receive if it is accepted, plus the probability it is rejected multiplied by the utility you would receive if your offer is rejected.

The number of terms in this expansion triples with each further step forwards– in english, the possibilities are as such, if we assume a constant utility when not proposing:

1. Accepted in round 0, control not flipped, accepted round 1
2. Accepted in round 0, control not flipped, rejected round 1
3. Accepted in round 0, control flipped, constant utility round 1
4. Rejected in round 0, control not flipped, accepted round 1
5. Rejected in round 0, control not flipped, rejected round 1
6. Rejected in round 0, control flipped, constant utility round 1

While this may get unwieldy, this is manageable for a few rounds, since all of the inputs here other than x are constants or can be assumed as constants, including the probability of acceptance (more on that in a moment).

From the Probabilistic Responder perspective, the formula is very simple, where x is the positive offer (and therefore the responder receives $-x$:

$$P(\text{Accept}(x)) = \frac{1}{1 + \text{sigmoid}(-\alpha \cdot (-x - u\_bad))}$$

Given that the responder receives x as a constant, all they can optimize is α. Therefore, we can simulate updates to strategies between the players to see if we cycle or reach a stable equilibrium. In the simulations, we used the Republicans in Control scenario for the proposer and responder utilities.

The first versions of the simulation involved simple versions of the two strategies. The Forward Looking Proposer optimized for the proposal value that maximized recursive expected utility for a given α, while the Probabilistic Responder optimized α for a given utility

[5] Code for this can be found in the repo inside the equilibrium.ipynb file

offer. This scenario reached equilibrium very quickly at a proposal of roughly 75 (the threshold) and a very low variance α. Adding in a rejection dependent bump to control flipping made no difference to the result. This result was quite obvious: the Responder didn't know that the proposer would update their proposal based on the chosen α, so had no incentive to be strategic, therefore moved as close to utilitarianism as was possible.

However, when we gave the Responder the ability to understand how the Forward Looking Proposer would respond to its α choice, suddenly we reached a new equilibrium that was much more favorable to the responder. By giving the responder the knowledge of how the proposer would respond to different α choices, the responder was now able to maximize its future value more efficiently, and therefore quickly selected a very high variance α value, which pushed the offer from 75 down to 42, a significant improvement.

As a final step, I attempted to then give the Proposer the knowledge that the Responder would pick the best α based on knowing how the proposer would respond to that chosen α value. I am unsure that I believe the result, although I don't see the bug in the code, but our outcome was a threshold proposal of 75 with a high variance response from the responder. I must admit that I don't see how this is better for the proposer than a lower variance approach.

Further work could be done here, too, but as a starting point, establishing that these realistic strategies play the game differently based on level-k strategies suggests a possibility of influencing the players, bringing us closer to our original goal.

5 Conclusions and Future Work

While this paper did not arrive at the lofty goal it set itself: creating a brand new system of democratic governance that protects norms, we nonetheless covered broad swaths of an under-explored space. We conclusively established that within our single shot framework punishments for shutdowns occurring did not change rational legislator behavior, examined various strategies parties might pursue in shutdown negotiations, and even went further in depth on two likely approaches.

However, this work is far from done. This project, initially envisioned as a D.Phil pursuit over years, was never going to be completed in two short months. At the same time, after half a decade sitting in the back of my mind at risk of never being started, has finally begun. Littered throughout the footnotes and asides of this paper are questions that could each merit fifteen pages of their own. After completing a course marrying my twin interests in computer science and game theory, I feel ready to take them on as well.