Learning in Multi-Issue Bargaining*

Renee Bowen[†] Ilwoo Hwang[‡] Stefan Krasa[§]

January 23, 2020

Abstract

How does learning about proposer power affect agents' ability to compromise? We study a dynamic multi-issue bargaining game between a proposer and a responder. Issues arrive at random, and with each new issue, the proposer makes a proposal, which the responder either accepts or rejects. In case of rejection, the proposer can attempt to force the issue and implement his ideal and will succeed with some probability that is a function of the proposer's unobserved ability. Both players learn about the proposer's ability over time as new issues arise. We show that there is conflict when the belief about the proposer is either high or low, but that compromise can occur in an intermediate region of beliefs. This is driven by both players' incentive to avoid learning in that region. We extend the model to include the possibility of difficult issues arising in which no compromise is possible. Difficult issues can disrupt a previously established compromise, forcing conflict for issues in which compromise was previously possible (easy issues). The reason conflict ensues is that the responder learns the proposer is weak and no longer has an incentive to compromise, even on easy issues.

*We thank the seminar audience at the Harris School of Public Policy, University of Chicago, and participants at the POLECONUK conference in Nottingham and the 2nd ETH Workshop on Democracy: Theoretical Political Economy, ETH Zurich for helpful comments and suggestions.

[†]UC San Diego and NBER; trbowen@ucsd.edu

[‡]University of Miami, Department of Economics, Coral Gables, FL 33124; ihwang@bus.miami.edu

[§]University of Illinois, Department of Economics, Urbana, IL 61801 USA; skrasa@illinois.edu

1 Introduction

The impact of agenda setters, such as a president, a prime minister, or a party leader in the process of political decision making has been the focus of research at least since the seminal contribution of Romer and Rosenthal (1979). In their paper, the power of the agenda setter is derived from a first-mover advantage. In this paper we investigate another source of power for an agenda setter, the ability to put pressure, cajole or lobby others to support a policy proposal. In other words, to *force the issue*. However, agenda setters may differ in their abilities to force an issue, and both responder and the agenda setter may only learn about the agenda setter's skills in the process of contentious bargaining battles.

The skill to convince is undoubtedly an important characteristic of a successful leader. For example, president Lincoln's ability to get the 13th amendment passed in the House, was due to his skills and active involvement in getting the necessary votes. More recently, President Clinton was able to establish good relationships with Republican legislators that helped him to get some of his agenda passed even after the Gingrich revolution. On the other hand, President Obama was often criticized for his frosty relationships with congress, which was at least in part responsible for the lack or passage of an immigration bill. Outside the political arena, there are examples of individuals in organizations who have to bargain over multiple issues over a period of time, with a new issue arising when one issue is closed. One such example is a dean in an academic institution making a proposal to faculty. At each meeting, a new curriculum change, hire, promotion, or other issue may arise and must be bargained over. Other examples include labor unions and firms or husband and wife.

In all these examples, some issues appear to be concluded relatively quickly with a compromise outcome, while for others conflict (possibly lengthy) may ensue with an uncertain outcome. In some situations, conflict may arise where it was previously not expected causing participants to wonder "why are we fighting about this?". We provide a first step to capture the dynamics of such repeated bargaining situations. The key assumption in the model is that upon rejection of a proposal, the proposer has an opportunity to *force the issue*. The power of the agenda setter is captured by his ability to force the issue (through pressure, deal-making or "going above the responder's head"). The ability of the proposer to force the issue may be unknown ex-ante to both the proposer and the responder, but with each new conflict, there is learning about the proposer's power. A *tough* proposer will succeed in forcing the issue with some probability, while a *soft* proposer will never be successful. We study a bargaining environment with learning about the proposer's bargaining ability and ask when we should expect to see conflict or compromise in bargaining.

To keep the analysis tractable we consider a stylized model in which each issue has three possible outcomes, the status quo, the proper's ideal, and a compromise proposal. The compromise outcome is the second most preferred for both players, and the responder prefers the status quo to the compromise. Time is continuous and a new issue arises with some probability in an interval of time. In the spirit of the dynamic bargaining literature (Rubinstein, 1982; Kalandrakis, 2004; Bowen and Zahran, 2012; Bowen et al., 2014; Dziuda and Loeper, 2016), we consider that there are multiple rounds in which a decision must be taken. Unlike much of the recent dynamic legislative bargaining literature we assume that in each round there is a new issue to decide (such as a new faculty to hire), so the status quo position of issues are not linked intertemporally. Once the issue arises, the proposer can immediately make a proposal. The responder then has a chance to accept or reject. If the responder rejects the proposal, then the proposer can choose to *force the issue*.

Learning about whether the proposer has the ability to successfully force the issue (or be tough) is modeled in the spirit of the exponential bandit literature as in Keller et al. (2005). There is a common prior probability that the proposer is tough. While the proposer is attempting to force the issue the common belief that the proposer is tough is decreasing as long as he has not successfully forced the issued. If the proposer succeeds in forcing the issue (by implementing his ideal), then players know that he is tough for certain and the belief jumps to one. New issues arise at random times so an issue can be abandoned before the proposer observes a success.

We show in this setting that there is a Markov perfect equilibrium in which players choose to compromise on issues for an intermediate range of beliefs. That is, compromise is possible if the proposer is neither too strong, nor too weak. As is standard in the bandit literature, the upper bound on the belief is driven by the proposers "exploitation" versus "exploration" tradeoff. The responder is known to accept the compromise when the proposer's type is high, and thus the compromise is endogenously a "safe" alternative. The safe alternative may be exploited, or the risky alternative (proposing his ideal) may be explored.

The lower bound on the belief is driven by the responder's incentives. When the proposer is sufficiently weak, the responder knows that the likelihood of the status quo staying in place is very high when the proposer attempts to force the issue. Thus the responder never accepts the compromise. The proposer, knowing that the compromise will be rejected, proposes his ideal because he has nothing to lose. He has some small probability of implementing his ideal and discovering that he is strong. When the belief that the proposer is tough is sufficiently high, the responder prefers not to risk the proposer's ideal being implemented and learning that the proposer is good.

This leads to the following dynamics. If at the beginning of the game the proposer is very strong (i.e., the belief that he is tough is high), then conflict will ensue over issues until the belief about the proposer's toughness drifts sufficiently low. When the belief becomes low enough it enters the compromise interval and the proposer chooses to compromise. If at the beginning of the game the proposer is very weak (i.e., belief that he is tough is low), then there will be perpetual conflict. The responder will never accept a compromise proposal knowing that the proposer has a very low probability of successfully forcing an issue. The proposer, knowing this, forces every issue, because he has nothing to lose – there is still a small probability that he is successful.

We extend the model to include the possibility that a "difficult" issue arises with some small probability. A difficult issue is one for which compromise is not an option. Either the proposer avoids conflict by proposing the status quo, or the proposer forces his ideal. It is assumed that difficult issues arise with sufficiently small probability. Otherwise, issues are as before and can take three positions – status quo, compromise and proposer's ideal. We now call these issues where compromise is possible "easy". We show in this extended model that two equilibria exist. In both equilibria compromise is implemented for an intermediate interval of beliefs when the issue is easy. For difficult issues, the strategies and outcomes differ qualitatively between the two equilibria.

We refer to the first equilibrium as the "conflict equilibrium". In the conflict equilibrium the proposer always proposes his ideal so there is always conflict (and learning) when a difficult issue arrives. Thus even if beliefs are in the interval where compromise is implemented for easy issues, in the long run they drift sufficiently low that conflict ensues even for easy issues. This helps to explain why conflict may arise for issues that may have once been resolvable. The reduction in proposer power leads the responder to no longer wish to compromise.

We call the second equilibrium "avoiding the issue". In this equilibrium, the proposer proposes his ideal when a difficult issue arises for all values of beliefs except the lower bound of the compromise interval. The dynamics in this equilibrium is quite different from the first. If beliefs are in the compromise interval, they will drift down with each difficult issue, but when they hit the compromise boundary there is no further conflict. One can interpret this as the responder "avoiding the issue". Once the boundary is hit there is no conflict on either difficult or easy issues.

Literature Review

There is an extensive literature on dynamic legislative bargaining. As in Diermeier and Fong (2011) (and Romer and Rosenthal (1979) in a static setting) we assume that there is a designated agenda setter. This assumption allows us to focus on learning about the type of only one of the players. In contrast, Baron and Ferejohn (1989) and Baron and Ferejohn (1987) consider the case of a decentralized committee in which each member can be selected to be the agenda setter. These models have been extended to multidimensional policy spaces by Banks and Duggan (2000) and Banks and Duggan (2006), and to sequential bargaining over different policies (c.f., Baron (1996), Kalandrakis (2004), Duggan and Kalandrakis (2012)). To link policies over time, these models assume that once a policy is enacted it determines the status quo point for the next period. In our model the status quo is exogenous and fixed, but the belief about the proposer's type links issues through time. This allows us to clearly identify the effects of learning in the bargaining process.

The choice between a risky proposal (the proposer's ideal) and a safe alternative (the compromise) is modeled as a bandit problem in the spirit of Keller et al. (2005). In this sense, this project is related to the growing literature on collective experimentation and voting rules, including Strulovici (2010), and Anesi and Bowen (ming). Like ours, these papers study the interaction between collective choice and experimentation, however, what is uncertain in our paper is the bargaining ability of the proposer. In equilibrium, the proposer's ideal action can be considered the risky alternative, while the compromise is the safe alternative. Interestingly, the proposer and responder have opposing incentives to experiment. This generates the possibility for an intermediate interval of beliefs such that the safe alternative is implemented. In this interval experimentation is too risky for the proposer, while not conveying enough information for the responder to trigger it. Other papers considering policy experimentation and collective choice include Majumdar and Mukand (2004), Volden et al. (2008), Cai and Treisman (2009), Callander (2011), Callander and Hummel (2014), Millner et al. (2014), Hirsch (2016) and Freer et al. (2018). Callander and Hummel (2014) consider a case in which a political party preemptively experiments on policy to affect future decisions of the opposition party. Our paper differs because in our model agents do not learn the type of the policy, but the strength of the proposer which endogenously determines the future outcome.

The existing literature on bargaining with incomplete information (Fudenberg et al., 1985; Abreu and Gul, 2000; Deneckere and Liang, 2006; Lee and Liu, 2013) typically focuses on the effect of private information of bargainer(s). In these models, a rejection by the informed bargainer signals that the bargainer has a higher reservation value. In contrast, in our paper players are symmetrically uninformed about the president's ability and their conflicts over policy induces social learning. Uncertainty about the bargaining strength of agents has been previously proposed as a rationale for delay in bargaining. This has been explored in the seminal works of Admati and Perry (1987), Cramton (1992) and more recently by Friedenberg (2019). We do not seek to explain delay in this paper, but rather we seek to explain when we expect compromise to arise, or when we expect a challenge to ensue. Our model features delay in the sense that conflict exogenously implies some delay relative to agreeing to a policy proposal immediately.

Theoretically our paper is also related to a rich literature in continuous time bargaining models, including Ortner (2019); Perry and Reny (1993). Ortner (2019) is closely related as, similar to us, the effect of evolving proposer power on bargaining outcomes is considered. Ortner (2019) considers a single issue, whereas we consider multiple. In addition, unlike Ortner (2019), our model features endogenous evolution of proposal power as conflict is chosen. The proposer and responder must therefore consider the trade-off between fighting for their preferred issue (which will imply learning about proposer strength) and settling for a less preferred outcome. In earlier work, Powell (2004) also takes up the question of endogenous evolution of proposer power, but in a discrete time setting with effectively a single issue being considered.

The remainder of the paper is organized as follows. In Section 2 we present the baseline dynamic bargaining model. In Section 3 we provide a benchmark result in which we fully characterize the unique equilibrium of the one-issue version of our model. In this model there is no role for learning and compromise arises because conflict involves delay. In Section 4 we characterize the equilibrium of the dynamic model with multiple issues and show that the set of beliefs for which compromise is possible in equilibrium shrinks. In Section 5 we extend the baseline model to include the possibility of more difficult issues and show that these can introduce learning that forces players into permanent conflict or permanent compromise.

2 Model

We present a stylized model of two players, a proposer P (for example a President) and a responder R (for example congress) who bargain over an infinite sequence of issues. We assume that time is continuous $t \in [0, \infty)$ and new issues arise at random times. In particular, in any interval of time $[t, t+dt)$ a new issue arrives with probability $1-e^{-\xi dt} \approx \xi dt$. At most one issue is bargained over at any time, so the arrival of a new issue means the previous issue is abandoned with no change in the status quo.

The game proceeds as follows. When a new issue arrives, the proposer can make a proposal $x \in X = \{x_0, x_c, x_P\}$. We denote x_0 as the exogenous status quo position, x_c as a compromise, and x_P represents the proposer's preferred option. When a proposal is made, the responder chooses to accept or reject the proposal. If the proposal is accepted, it is implemented immediately and the proposer and responder receive payoffs. If the proposal is rejected, the proposer can choose to *force the issue* and implement his ideal with some probability or allow the status quo to be implemented immediately. We assume that the proposal, the decision to accept/reject and the decision to force the issue do not introduce costly delay. In other words we assume that they happen instantaneously.

We model the proposer's choice to force the issue or not as the choice of a two-armed exponential bandit in the spirit of Keller et al. (2005). If the proposer chooses to force the issue his success at implementing his ideal x_P depends on his type $\theta \in \{0, 1\}$. We say the proposer is *tough* or $\theta = 1$ if he is successful at forcing an issue with probability $1 - e^{-\lambda dt} \approx \lambda dt$ over time interval $[t, t + dt)$. If the proposer is *soft* then he is never successful at forcing an issue. The common prior probability that the proposer is tough is p_0. We say the proposer is *weak* if p is low and *strong* when p is high. If the proposer chooses to force the issue, the issue is resolved either when the proposer is successful and x is implemented, or if the proposer subsequently makes an offer which is accepted.[1]

Players receive payoffs $u_P(x)$ and $u_R(x)$ when position x is implemented. We assume that R strictly prefers x_0 to x_P, and, similarly, P strictly prefers x_P to x_0. We normalize the utility of players such that $u_P(x_0) = u_R(x_P) = 0$. Also, let $u_P(x_P) = \bar{u}$, $u_P(x_c) = u_C$, $u_R(x_0) = \bar{v}$ and $u_R(x_c) = v_C$. We assume that $0 < u_C < \bar{u}$, and $0 < v_C < \bar{v}$.[2] Note that x_c is considered a compromise proposal because it is the second-ranked outcome for both players. Note also that in this setting $v_C + u_C > \max\{\bar{u}, \bar{v}\}$ is sufficient for compromise to be efficient.[3] Finally, utility is discounted at a rate e^{-rdt} with $r > 0$ and players maximize discounted sums of payoffs from all issues bargained over.

Learning If the responder accepts the proposer's offer, or the proposer decides not to force the issue, then there is no learning about the proposer's type and beliefs are unchanged. If the responder rejects the offer and the proposer attempts to force the issue, the belief changes depending on the outcome of the attempt. If the proposer succeeds in forcing the issue, then the belief that the proposer is tough jumps to one and there is no learning thereafter. In other words, "good" news is conclusive. As the proposer attempts to force the issue without success, then players become more pessimistic about the proposer's toughness. Formally, while the proposer is forcing the issue, the belief that the proposer is tough changes on the time interval $[t, t + dt)$

[1] The way we model "forcing the issue" has parallels to Lee and Liu (2013). They also assume the proposer has two types and a good proposer has higher probability of successfully extracting higher outside payment in the event of no disagreement. Lee and Liu (2013) focus on one-sided learning, whereas we focus on two-sided learning in a simpler setting.

[2] Note that commonly used utility functions satisfy these minimal assumptions. For example $u_i(x) = -(x - x_i)^2$ with $x \in \mathbb{R}$, $x_0 < x_P$, $x_R = x_0$, $x_C \in (x_0, x_P)$.

[3] The necessary and sufficient condition for compromise to be always efficient is $u_C + v_C > \max\{(\lambda\bar{u} + \xi\bar{v})/(r + \xi + \lambda), \xi\bar{v}/(r + \xi)\}$. This is because of the loss due to delay when there is conflict.

via Bayes' rule according to

$$p_{t+dt} = \frac{p_t(1 - \lambda dt)}{p_t(1 - \lambda dt) + (1 - p_t)}. \quad (1)$$

Markov Strategies We restrict attention to stationary Markov strategies where the state of the game at period t is given by the belief p_t about the probability that the proposer is tough. Denote an offer strategy for the proposer as $\chi : [0,1] \to X$ which maps a belief p_t into a proposal in X. An acceptance strategy for the responder is a correspondence $A : [0,1] \rightrightarrows X$ that gives the set of proposals which the responder will accept given the state p_t. Finally the strategy of the proposer to force the issue or not is given by $\beta : [0,1] \times X \to \{0,1\}$, where 1 indicates that the issue is forced.

To ensure that strategies are well-defined in this continuous time setting, we assume *admissability* in the sense of Klein and Rady (2011).[4] We further restrict attention to strategies such that the responder accepts proposals when indifferent.[5] This ensures that the equilibrium of the continuous-time game is the limit of an equilibrium of a corresponding discrete-time game. These restrictions together imply that equilibrium acceptance sets $A(p)$ will be closed. We consider Markov perfect equilibria which are subgame perfect equilibria in which players use Markov strategies. We refer to a Markov perfect equilibrium with the above restrictions as simply an *equilibrium*.

3 Single issue

To build intuition for the model and provide a benchmark where learning plays no role we first describe the outcome of a game in which only one issue is bargained over. In this game, time is continuous. At time $t = 0$ the proposer chooses a proposal $x \in \{x_0, x_c, x_P\}$, the responder then decides to accept or reject, and, finally, if rejected, the proposer decides whether or not to force the issue. All these decisions happen instantaneously. If the proposer decides to force the issue, either it is successful at rate λ, in which case the game ends with the proposer's ideal

[4]Following Klein and Rady (2011) the strategy profile $\{\chi, A, \beta\}$ is *admissable* if there exists at least one well-defined solution to the corresponding law of motion for posterior beliefs. This is the case if and only if for each initial belief p_0, there is a function $t \mapsto p_t$ on $[0, \infty)$ that satisfies

$$p_t = \frac{p_0 e^{-\lambda \int_0^t I_{\chi(p_\tau) \notin A(p_\tau)} I_{\beta(p_\tau, \chi(p_\tau))=1} d\tau}}{p_0 e^{-\lambda \int_0^t I_{\chi(p_\tau) \notin A(p_\tau)} I_{\beta(p_\tau, \chi(p_\tau))=1} d\tau} + 1 - p_0}. \quad (2)$$

[5]The is a common assumption in the dynamic bargaining literature, for example, Baron and Ferejohn (1989) or Bowen et al. (2014).

implemented, or the issue is abandoned at rate ξ after which the game ends and the status quo is implemented. The decision to fore the issue is irreversible.[6]

We solve by backward induction beginning with the decision to force the issue. If a proposal x has been rejected, then the responder decides to force the issue or allow the status quo x_0 to be implemented. Forcing the issue results in the proposer's ideal x_P being implemented with some probability, hence it is always optimal to force the issue regardless of the state. Thus $\beta(p) = 1$ for all p.

Now consider the responder's incentive to accept or reject the proposal. To do this we must calculate the players' dynamic payoffs under rejection given that rejection leads to the issue being forced. While the issue is being forced, in the time interval $[t, t + dt)$ the probability that the proposer succeeds is given by $1 - e^{-p\lambda dt} \approx p\lambda dt$. In this case, the game ends, and the proposer and responder receive payoffs $\bar{u}$ and 0 respectively. The probability that the issue is abandoned in the interval $[t, t + dt)$ is $1 - e^{-\xi dt} \approx \xi dt$. This is independent of whether or not the issue has been successfully forced. If the issue is abandoned, then the proposer has lost his opportunity to successfully force the issue and the game ends with the status quo implemented. The payoffs to the proposer and responder are thus 0 and $\bar{v}$ respectively. With the remaining probability $1 - p\lambda dt - \xi dt$ nothing happens in that time interval and players continue to wait either for the issue to be successfully forced or for the issue to be abandoned.

Denote the dynamic payoff for player i when the proposal is rejected as $V_{i,reject}$. This is determined recursively by the equation

$$V_{i,reject} = p\lambda dt u_i(x_p) + \xi dt u_i(x_0) + (1 - p\lambda dt - \xi dt)e^{-rdt} V_{i,reject}.$$

Substituting $1 - rdt$ for e^{-rdt}, dropping higher order terms and simplifying yields

$$V_{i,reject} = \frac{p\lambda u_i(x_p) + \xi u_i(x_0)}{p\lambda + \xi + r}.$$

The responder thus accepts the proposal if $u_R(x) \geq \xi\bar{v}/(p\lambda + \xi + r)$. The responder's payoff to the proposer's ideal is normalized to 0 so this is always rejected. The responder's payoff to the status quo is $\bar{v}$ so this is always accepted. The responder accepts the compromise if and only if $v_C/\bar{v} \geq \xi/(p\lambda + \xi + r)$ or if

$$p \geq \frac{\xi(\bar{v} - v_C) - rv_C}{\lambda v_C} \equiv \underline{p}_1.$$

Consider the proposer's choice of proposal $x \in \{x_0, x_c, x_p\}$. The proposer's ideal x_p is always rejected, and the status quo x_0 is always accepted. Given that the proposer prefers some chance

[6] See the discussion at the end of this section considering if this is relaxed and compromise can be selected after the proposer chooses to force the issue. The results in this case are unchanged.

of achieving his ideal payoff, proposing his ideal always dominates proposing the status quo. We now compare the proposer's incentive to propose the compromise versus his ideal. First consider that $v_C/\overline{v} < \xi/(p\lambda+\xi+r)$. Then the responder rejects the compromise and the proposer is better of proposing his ideal with some probability of implementation. Next suppose $v_C/\overline{v} \geq \xi/(p\lambda+\xi+r)$, then if the proposer proposes the compromise it is accepted right away, but his ideal is implemented with some probability and with some delay. The proposer will propose the compromise if and only if $u_C/\bar{u} \geq \frac{p\lambda}{p\lambda+\xi+r}$, or if and only if

$$p \leq \frac{u_C(r+\xi)}{\lambda(\overline{u}-u_C)} \equiv \overline{p}_1. \tag{3}$$

We require $\underline{p}_1 < \overline{p}_1$ for compromise to occur which implies $v_C/\overline{v} > [\xi/(r+\xi)][1-u_C/\bar{u}]$. We therefore have the following result in the case of one issue.

Proposition 1 *In the one-issue model, the proposer always forces the issue after a rejection. i.e.* $\beta(p,x) = 1$ *for all* p *and all* x*. In addition:*

1. *Suppose* $v_C/\overline{v} \leq [\xi/(r+\xi)][1-u_C/\bar{u}]$*, then there is always conflict in the one issue case. Specifically,*

$$\begin{aligned} \chi(p) &= x_p, \\ A(p) &= \{x_0\}. \end{aligned}$$

2. *Suppose* $v_C/\overline{v} > [\xi/(r+\xi)][1-u_C/\bar{u}]$*, then there is compromise if the proposer is uncertain about his strength. Specifically,*

$$\chi(p) = \begin{cases} x_c & \text{if } p \in [\underline{p}_1, \overline{p}_1] \\ x_p & \text{otherwise.} \end{cases}$$

$$A(p) = \begin{cases} \{x_0, x_c\} & \text{if } p \in [\underline{p}_1, 1] \\ \{x_0\} & \text{otherwise.} \end{cases}$$

Note from equation (3) that $\overline{p}_1 > 0$ for all parameter values. Corollary 1 summarizes the parameters such that compromise is possible.

Corollary 1 *In the one-issue model compromise can occur in equilibrium if and only if one of the following parameter restrictions hold:*

1. $\underline{p}_1 < 0 < 1 < \overline{p}_1$*, in which case* $u_C/\overline{u} > \lambda/(\lambda+\xi+r)$*,* $v_C/\bar{v} > \xi/(r+\xi)$*;*

2. $\underline{p}_1 < 0 < \overline{p}_1 \leq 1$, *in which case* $u_C/\overline{u} \leq \lambda/(\lambda + \xi + r)$, $v_C/\overline{v} > \xi/(r + \xi)$.

3. $0 \leq \underline{p}_1 < 1 < \overline{p}_1$, *in which case* $u_C/\overline{u} > \lambda/(\lambda + \xi + r)$, $\xi/(\lambda + r + \xi) < v_C/\overline{v} \leq \xi/(r + \xi)$;

4. $0 \leq \underline{p}_1 < \overline{p}_1 \leq 1$, *in which case* $u_C/\overline{u} \leq \lambda/(\lambda+\xi+r)$, $[\xi/(r+\xi)][1-u_C/\overline{u}] < v_C/\overline{v} \leq \xi/(r+\xi)$;

We illustrate the parameters such that compromise can occur in Figure 1. Note that compromise occurs in the one issue model because of the cost of delay induced by conflict and because the proposer prefers compromise to the probability of a worse outcome. This happens only when the responder's payoff to compromise is sufficiently high.[7] In the next section we explore the role of learning in the fully dynamic model with multiple issues.

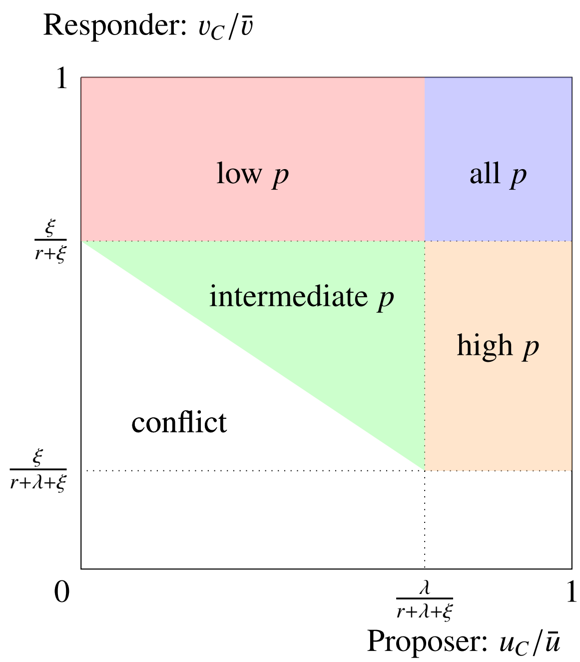

Figure 1: One issue range of parameter values such that compromise occurs in equilibrium. Blue region is for always compromise $\underline{p}_1 < 0 < 1 < \overline{p}_1$, red is for compromise when p is sufficiently low $\underline{p}_1 < 0 < \overline{p}_1 \leq 1$, orange is for p sufficiently high $0 \leq \underline{p}_1 < 1 < \overline{p}_1$, and green is for p in an intermediate range $0 \leq \underline{p}_1 < \overline{p}_1 \leq 1$.

Note that the case considered in this section involved no learning, but is equivalent to the following single-issue case with learning. Consider a single-issue case, but assume that there is learning and the proposer is able to switch to compromise at any time after he has decided to force the issue. The bounds on the interval of beliefs such that compromise occurs is identical

[7] In the Appendix Section 7.4 we present a version of the model in which there is no benefit to delay from the compromise, and we attain qualitatively similar results for the case of learning.

to those derived in this section. To see this, note that for the proposer to propose the compromise, he must be indifferent between continuing to force the issue, and immediate compromise. Similarly, the responder must be indifferent between immediate compromise and the continued possibility of the proposer's ideal. Thus the threshold beliefs satisfy the same indifference conditions as before. One way to interpret this is that learning has no value when there are no further issues with which to make use of the knowledge.

4 Learning over multiple issues

We study Markov perfect equilibria of the game with learning about the proposer's type over an infinite sequence of issues. The payoff relevant state variable of the game is the belief that the proposer is tough p. The main insight is that learning reduces opportunities to compromise. We first characterize the Markov perfect equilibrium of the game with compromise. Note that we do not present the proposer's strategy to force the issue or not as forcing the issue occurs trivially in every equilibrium as the intuition in the one-issue case of Section 3 points to.

Proposition 2 *There exists an equilibrium when there are multiple issues as follows.*

$$\chi(p) = \begin{cases} x_c & \text{if } p \in [\underline{p}, \overline{p}] \\ x_P & \text{otherwise,} \end{cases} \qquad A(p) = \begin{cases} \{x_0, x_c\} & \text{if } p \geq \underline{p} \\ \{x_0\} & \text{if } p < \underline{p}, \end{cases}$$

where

$$\begin{aligned} \overline{p} &= \frac{u_C r(\lambda + r + \xi)}{[(\bar{u} - u_C)(r + \lambda) - \xi u_C]\lambda} \\ \underline{p} &= \frac{[\xi(\bar{v} - v_C) - r v_C](r + \lambda + \xi)}{\xi \bar{v} \lambda}. \end{aligned}$$

We provide the complete proof in Appendix Section 7.1, and henceforth all proofs are included in the Appendix unless otherwise stated. We summarize the key incentive constraints to provide intuition for the result.

Permanent Conflict We first derive the value function that would arise if there is permanent conflict, i.e., the proposer offers x_P in every period, which the responder always rejects. This occurs for beliefs $p < \underline{p}$. If the proposer does not successfully force the issue, the belief p gradually decreases, and thus stays in the conflict region. If the issue is successfully forced, then the belief jumps up to one, and thus conflict continues.

In the dynamic model an issue is resolved if either a proposal is accepted, or the issue was successfully forced. In either case, players are simply waiting for a new issue to arrive and there is no learning. Define $\hat{V}_{i,c}$ to be the player i's value function if the issue has been resolved. In the following we consider an infinitesimally small time interval $[t, t+dt)$. First, note that if the current issue has been resolved, then the players do not take any action and wait for the next issue to come up, which occurs with probability ξdt. Then for $i = P, R$ and $p \in [0,1]$,

$$\hat{V}_{i,c}(p) = e^{-rdt}(\xi dt V_{i,c}(p) + (1-\xi dt)\hat{V}_{i,c}(p)).$$

Rearranging terms, noting that $e^{-rdt} \approx 1 - rdt$, dropping higher order terms and simplifying yields

$$\hat{V}_{i,c}(p) = \frac{\xi}{r+\xi} V_{i,c}(p). \tag{4}$$

Now consider that the proposer has decided to force the issue and is waiting for a success or for a new issue to arrive. We denote these value functions by $V_{P,c}(p)$ and $V_{R,c}(p)$, respectively. Using $\hat{V}_{i,c}(p)$ we derive $V_{i,c}(p)$ for the two extreme values of p. Suppose first that the proposer is known to be tough, i.e., $p = 1$. There is a conflict at every instant, and the proposer succeeds in forcing the issue at a rate of λ, in which case the proposer and the responder obtain payoffs $\bar{u}$ and 0, respectively, and the players wait for the next issue to come up. If the issue has not been resolved, then with probability ξdt the issue is replaced with the new one, in which case P and R obtain payoffs of 0 and $\bar{v}$, respectively. Therefore,

$$\begin{aligned} V_{P,c}(1) &= \lambda dt[\bar{u} + \hat{V}_{P,c}(1)] + (1-\lambda dt)e^{-rdt}V_{P,c}(1), \\ V_{R,c}(1) &= \lambda dt\hat{V}_{R,c}(1) + (1-\lambda dt)(\xi dt\bar{v} + e^{-rdt}V_{R,c}(1)). \end{aligned}$$

Simplifying, we have

$$\begin{aligned} V_{P,c}(1) &= \frac{\lambda(r+\xi)}{r(r+\lambda+\xi)}\bar{u}, \\ V_{R,c}(1) &= \frac{\xi(r+\xi)}{r(r+\lambda+\xi)}\bar{v}. \end{aligned}$$

When $p = 0$, the proposer never succeeds in forcing the issue (that is, $\lambda = 0$ in the above equations), and thus the value functions are

$$\begin{aligned} V_{P,c}(0) &= 0, \\ V_{R,c}(0) &= \frac{\xi}{r}\bar{v}. \end{aligned}$$

Player i's value function $V_{i,c}(p)$ is the convex combination of $V_{i,c}(1)$ and $V_{i,c}(0)$. To understand this, note that the players' never change their actions in the future: Regardless of the outcome,

P always proposes x_P which R rejects. Therefore,

$$V_{P,c}(p) = pV_{P,c}(1) + (1-p)V_{P,c}(0) = p\bar{u}\frac{\lambda(r+\xi)}{r(r+\lambda+\xi)}, \tag{5}$$

$$V_{R,c}(p) = pV_{R,c}(1) + (1-p)V_{R,c}(0) = \frac{\xi\bar{v}}{r}\left(1 - \frac{p\lambda}{r+\lambda+\xi}\right). \tag{6}$$

Note that both value functions are affine linear. The proposer's function is increasing in p, while the responder's is decreasing in p.

No Conflict The second important benchmark is where no conflict occurs or when $\underline{p} \leq p \leq \overline{p}$. In this case, the proposer offers the compromise x_c, and the responder accepts. No learning occurs, and hence the payoffs do not depend on the belief p. Let $V_{i,n}$ be player i's continuation utility. We can again define $\hat{V}_{i,n}$ as the continuation utility after the issue has been resolved. The same argument used to derive (4) can be used to show that

$$\hat{V}_{i,n} = \frac{\xi}{r+\xi}V_{i,n}.$$

Then

$$V_{P,n} = u_C + e^{-rdt}(\xi dt V_{P,n} + (1-\xi dt)\hat{V}_{P,n}),$$
$$V_{R,n} = v_C + e^{-rdt}(\xi dt V_{R,n} + (1-\xi dt)\hat{V}_{R,n}).$$

Simplifying yields

$$V_{P,n} = \frac{r+\xi}{r}u_C, \tag{7}$$

$$V_{R,n} = \frac{r+\xi}{r}v_C. \tag{8}$$

Upper bound on p We can find $\bar{p}$ as it is the p such that the proposer is just indifferent between conflict and compromise. The value matching condition (Dixit (2002)) gives

$$\frac{r+\xi}{r}u_C = \lambda p dt[\bar{u} + \hat{V}_{P,c}(1)] + (1-\lambda p dt)e^{-rdt}\frac{r+\xi}{r}u_C.$$

Simplifying yields

$$\bar{p} = \frac{u_C r(\lambda + r + \xi)}{[(\bar{u} - u_C)(r+\lambda) - \xi u_C]\lambda}.$$

Lower bound on p The lower bound on p is such that the responder is just indifferent between compromise and permanent conflict and this gives

$$\frac{r+\xi}{r}v_C = \frac{\xi\bar{v}}{r}\left(1 - \frac{p\lambda}{r+\lambda+\xi}\right).$$

This implies

$$\underline{p} = \frac{[\xi\bar{v} - (\xi + r)v_C](r+\lambda+\xi)}{\xi\bar{v}\lambda}.$$

We can compare the bounds on p such that compromise is achieved in the case of multiple issues to those derived in the single-issue case with no learning, and show that learning with multiple issues reduces the values of p such that compromise is attained.

Proposition 3 *When compromise is attained for an intermediate set of beliefs (i.e.* $0 < \underline{p} < \overline{p} < 1$*), the set of beliefs such that compromise is the outcome in the multi-issue case is smaller than the set of beliefs in the single-issue case. i.e.* $[\underline{p}, \overline{p}] \subset [\underline{p}_1, \overline{p}_1]$.

Note that the key difference is learning combined with multiple issues. Consider an alternate multi-issue model with no learning. That is, the probability of the proposer successfully forcing the issue is fixed at $p\lambda$, and thus the proposer decides at the beginning of the game whether to pursue permanent conflict or compromise. Multiple issues may arise, but these will not change the proposer's or responder's decisions. This therefore results in a scaling up of the payoffs both to conflict and compromise, relative to the single-issue case, and the bounds are thus unchanged. This highlights that the key driver of the reduction of the belief interval going from single-issue to multiple issues is learning combined with the future issues.

Thus learning with multiple issues reduces the opportunity to compromise. The intuition for the upper bound is a straightforward application of the standard dynamic experimentation versus exploitation trade-off in the bandit literature. From the perspective of the proposer, forcing the issue is risky, but the incentive to learn and benefit from that learning increases the incentive to take the risk. The proposer thus "experiments" with forcing the issue for lower values of the belief than if there was no learning possible.

The intuition for the lower bound is more subtle as it increases going from the case of no learning to learning and is driven by the responder's constraint. Like the proposer, the responder has a stronger incentive to experiment when learning is present and there are multiple-issues, however the responder hopes to learn that the proposer is weak. Since the responder prefers to experiment for low values of p (as opposed to high-values for the responder) an increase in the incentive to experiment raises the lower bound on the compromise interval.

In summary, learning reduces the set of beliefs for which compromise occurs. However, this does not imply that compromise is achieved less frequently under learning. In particular, suppose we start at a belief $p > \overline{p}_1$. Without learning we have permanent conflict. With learning, the beliefs move into the compromise region with a probability of at least $1 - p$, and hence eventual compromise is possible.

Value of information To provide further intuition for the result, we use the value functions to determine the value of information. The value of information is simply the expected informational benefit from conflict.[8] For the proposer, the benefit is in learning that he is tough, and for the responder, the benefit is learning that the proposer is soft. The value of information for player $i = P, R$ in the interval $[t, t + dt)$ is given by

$$\mathrm{VI}_i(p) = \lambda p(V_i(1) - V_i(p)) - V_i'(p)\lambda p(1 - p). \tag{9}$$

The first term in the expression is the expected benefit from learning that the proposer is tough for certain. The second term is the loss to the proposer if no success occurs, or the benefit to the responder if no success occurs. Figure 2 graphs the value of information and the value functions for a parametric example.

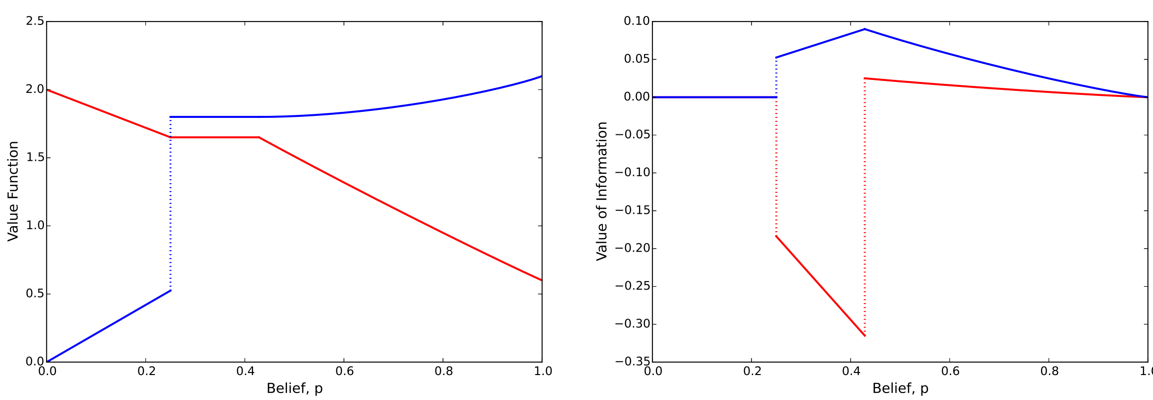

Figure 2: Utility and the value of information for the responder (red) and the proposer (blue): $\xi = 0.2$, $r = 0.1$, $\lambda = 0.7$, $v_C = 0.55$, $u_C = 0.6$, $\bar{v} = \bar{u} = 1$.

For $p < \underline{p}$ we have $V_i(p) = V_{i,c}(p)$ from equations (5) and (6). Simplifying (9) shows that the value of information is zero as illustrated in Figure 2. The reason is that information will not change players' strategies, as conflict will ensue even if there is new information.

[8]This is similar to the value of playing the risky alternative in Keller et al. (2005).

If $p \in [\underline{p}, \overline{p}]$ then $V_i(p) = V_{i,n}$ from equations (7) and (8), which is constant and hence $V_i'(p) = 0$. For the responder $V_R(1) < V_R(p)$ if and only if $v_C/\bar{v} > \xi/(r + \lambda + \xi)$. For the proposer, $V_P(1) > V_P(p)$ if and only if $u_C/\bar{u} < \lambda/(r + \lambda + \xi)$, which is the condition under which $\overline{p} < 1$. In this case the value of information is strictly positive and strictly increasing for the proposer, and strictly negative and strictly decreasing for the responder in the compromise region. This reflects the fact that as p gets smaller, the proposer's value of experimenting is decreasing, as there is less the proposer is able to do. From the responder's perspective, his benefit from information is increasing as beliefs drift down, because he approaches the region of beliefs such that he rejects the compromise.

In the Appendix, Section 7.1 we show that if $p > \overline{p}$ then the proposer's and responder's value functions are convex. If we start with some belief $p > \overline{p}$ at time t and learning occurs, then the belief at $t + dt$ remains above $\overline{p}$. By (9), the strict convexity of $V_i(p)$ implies that the value of information is strictly positive for both players. That is, both players have the ability to modify actions based on information arrival. Note that the proposer's value of information is continuous at $\overline{p}$, but the responder's is discontinuous at that point. The reason is that it is the proposer who chooses to switch from proposing conflict, to compromise. When the proposer does this the responder's value of information becomes negative—information can only hurt the responder by causing the proposer to revert to conflict.

We next examine parameter values that admit compromise in the multi-issue case and also show that this set is smaller than in the case of a single-issue, consistent with compromise being possible in fewer environments when there is multi-issue bargaining.

Regions of compromise The following corollary provides conditions under which there exists an equilibrium with a non-empty compromise interval.

Corollary 2 *In the multi-issue model compromise occurs if and only if one of the following parameter restrictions hold:*

1. $\underline{p} < 0 < 1 < \overline{p}$, *in which case* $v_C/\bar{v} > \xi/(r + \xi)$, *and* $u_C/\bar{u} > \lambda/(r + \lambda + \xi)$.

2. $\underline{p} < 0 < \overline{p} \leq 1$, *in which case* $v_C/\bar{v} \geq \xi/(r + \xi)$, *and* $u_C/\bar{u} \leq \lambda/(r + \lambda + \xi)$.

3. $0 \leq \underline{p} < 1 < \overline{p}$, *in which case* $u_C/\bar{u} > \lambda/(r + \lambda + \xi)$, $\xi/(r + \lambda + \xi) < v_C/\bar{v} \leq \xi/(r + \xi)$.

4. $0 \leq \underline{p} < \overline{p} \leq 1$, *in which case* $u_C/\bar{u} \leq \lambda/(r + \lambda + \xi)$ *and* $\xi/(r + \xi)\left[1 - \frac{r u_C}{(r+\lambda)\bar{u} - (r+\lambda+\xi)u_C}\right] < v_C/\bar{v} \leq \xi/(r + \xi)$.

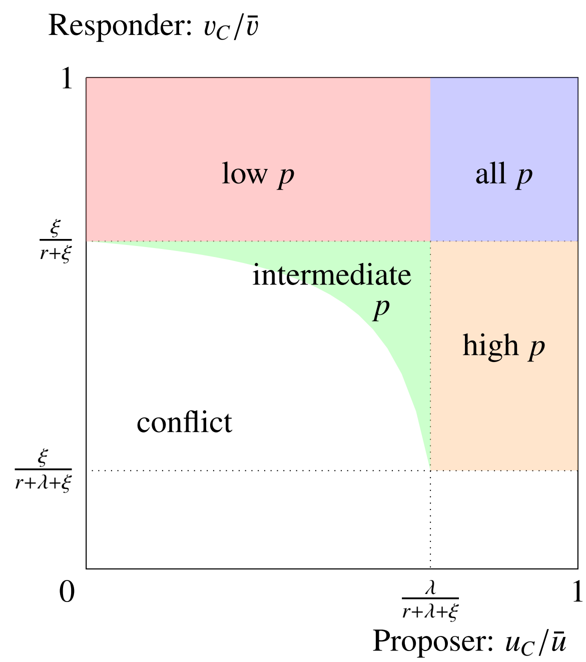

Figure 3: Multi-issue range of parameter values such that compromise occurs in equilibrium. Blue region is for always compromise $\underline{p}_1 < 0 < 1 < \overline{p}_1$, red is for compromise when p is sufficiently low $\underline{p}_1 < 0 < \overline{p}_1 \leq 1$, orange is for p sufficiently high $0 \leq \underline{p}_1 < 1 < \overline{p}_1$, and green is for p in an intermediate range $0 \leq \underline{p}_1 < \overline{p}_1 \leq 1$.

Finally, in all remaining cases there is always conflict. Figure 3 indicates the values of $v_C/\bar{v}$ and $u_C/\bar{u}$ such that these cases occur.

Note that the curved line is the lower bound of the region for which compromise occurs in a middle interval. Comparing Figures 1 and 3 it is clear that learning reduces the set of parameters that admit an equilibrium that will feature compromise. This can be seen by comparing condition 4 in Corollaries 1 and 2. In the single issue case the restriction on the responder's payoff is $[\xi/(r+\xi)][1 - u_C/\bar{u}] < v_C/\bar{v} < \xi/(r+\xi)$ and in the multi-issue case it is $[\xi/(r+\xi)][1 - \frac{ru_C}{(r+\lambda)\bar{u}-(r+\lambda+\xi)u_C}] < v_C/\bar{v} < \xi/(r+\xi)$. We can show that the lower bound in the multi-issue case is strictly above the upper bound in the single-issue case if $u_C/\bar{u} < \lambda/(r+\lambda+\xi)$ which is always the case in the green region.

Comparative Statics Given the simple expressions for the bounds on beliefs for compromise to occur, we can do straightforward comparative statics.

Proposition 4 *Suppose that* $0 < \underline{p} < \overline{p} < 1$.

1. *The set of beliefs such that compromise occurs is strictly increasing in r. i.e.*

$$\partial \underline{p}/\partial r < 0 \textit{ and } \partial \overline{p}/\partial r > 0$$

.

2. *The set of beliefs such that compromise occurs is strictly decreasing in λ. i.e.*

$$\partial \underline{p}/\partial \lambda > 0 \textit{ and } \partial \overline{p}/\partial \lambda < 0.$$

The first part of Proposition 4 is straightforward and says that compromise will occur in more environments when players are less patient. This is driven by the fact that a compromise is resolved quickly, but delay is introduced with conflict. The second part of Proposition 4 relates to the speed of learning. As the speed of learning increases (an increase in λ) players' willingness to compromise decreases. From (9) observe that the value of information increases with the speed of learning λ. As the value of information increases, the incentive to experiment increases, and thus the incentive to compromise decreases.

The results of this section suggest that there may be a benefit to turnover in proposers. If bargaining is with a new proposer each time, so learning has no value, then conflict is likely to be reduced. We next explore another possibility, that there are more *types* of issues possible.

5 Difficult issues

In this section, we assume that a "difficult" policy issue may arise. Formally, with probability $\alpha > 0$, a policy issue with no compromise proposal arises, and the proposer can only propose either x_P or x_0. With the complementary probability, the standard policy issue (with three possible outcomes) arises. We refer to the standard policy issue as an "easy" issue.

We focus on the parameter range such that compromise occurs in the intermediate range of beliefs. We assume that α is small, and thus the equilibrium behavior under the standard policy issue remains qualitatively the same. The compromise region is a closed interval $[\underline{p}_\alpha, \overline{p}_\alpha]$.

We construct two types of equilibria whose behavior differ qualitatively under difficult policy issues. First, there exists an equilibrium in which the proposer offers x_P for any $p \in (0, 1)$. We call this the *conflict equilibrium*. Second, there exists an equilibrium in which the proposer offers x_0 at $p = \underline{p}_\alpha$, and offers x_P otherwise. We call this the *avoiding the issue* equilibrium. We find the parametric conditions for each equilibrium to exist.

Proposition 5 (conflict equilibrium) *Let* $\underline{p}$ *be as defined in* (3). *Suppose that* α *is sufficiently small. Also, suppose that*

$$\frac{\xi(1-\alpha)}{r+\xi}\frac{u_C}{\bar{u}}+\frac{r+\xi}{\xi}\frac{v_C}{\bar{v}}\leq 1. \tag{10}$$

Then there exists an equilibrium with the following behavior:

- *Under the standard policy issues, the players compromise at* x_c *if and only if* $p \in [\underline{p}, \overline{p}_\alpha]$ *for some* $\overline{p}_\alpha \in (\underline{p}, 1)$.
- *Under the difficult policy issues, for any* $p \in (0,1)$ *the proposer offers* x_P *and the responder rejects the proposer's offer.*

In the *conflict equilibrium*, even if compromise can be maintained on easy issues, difficult issues arise. Conflict over difficult issues leads to learning. If the belief jumps to one, or if the belief drifts sufficiently low then it enters the conflict region even for easy issues. There is permanent conflict thereafter in equilibrium. This helps to explain why conflict can arise on seemingly easy, non-contentious issues, where there was no conflict before. The reason is that learning about proposer power generates bargaining dynamics that give either the proposer or the responder greater incentive to reject compromise proposals.

Proposition 6 (avoiding the issue) *Define* $\underline{p}_0$ *and* $\tilde{p}$ *as*

$$\begin{aligned}\underline{p}_0 &= \frac{(r+\lambda+\xi)(\xi(1-\alpha)\bar{v}-(r+\xi(1-\alpha))v_C)}{\lambda\xi\bar{v}}\\ \tilde{p} &= \frac{u_C}{\bar{u}}\frac{\xi(1-\alpha)(r+\lambda+\xi)}{\lambda(r+\xi)}.\end{aligned}$$

Suppose that α *is sufficiently small and*

$$\frac{(1-\alpha)(\bar{u}(r+\xi)-\xi u_C)\xi}{\bar{u}(r+\xi)(r+\xi(1-\alpha))} \leq \frac{v_C}{\bar{v}}, \tag{11}$$

then there exists a $\hat{p} \in [\underline{p}_0, \tilde{p}]$*, such that for any* $\underline{p}_\alpha \in [\underline{p}_0, \hat{p}]$*, there is an equilibrium with the following behavior:*

- *Under the standard policy issues, the players compromise at* x_c *if and only if* $p \in [\underline{p}_\alpha, \overline{p}_\alpha]$ *for some* $\overline{p}_\alpha \in (\underline{p}_\alpha, 1)$.
- *Under difficult policy issues, at* $p = \underline{p}_\alpha$ *the proposer offers* x_0 *and the responder accepts the proposer's offer. Otherwise, the proposer offers* x_P *and the responder rejects the offer.*

Moreover if

$$\frac{\xi(1-\alpha)}{r+\xi}\frac{u_C}{\bar{u}} + \frac{r+\xi}{\xi}\frac{v_C}{\bar{v}} \;\leq\; 1, \tag{12}$$

then $\tilde{p} \leq \underline{p}$.

In the *avoiding the issue* equilibrium, conflict ensues on difficult issues until the belief equals the lower bound of the compromise interval. At the lower bound the proposer accommodates on difficult issues by proposing the status quo. This avoids learning. There is compromise thereafter, even on difficult issues. Note that the lower bound on beliefs such that compromise will arise $\underline{p}_0$ is lower than the lower bound for the conflict equilibrium, but the upperbound is the same $\overline{p}_\alpha$. Thus the avoiding the issue equilibrium admits compromise on easy issues for a larger set of parameters than the conflict equilibrium.

Note that there is a set of parameters such that both equilibria exist. Specifically, conditions (10) and (11) together imply

$$\frac{(1-\alpha)(\bar{u}(r+\xi)-\xi u_C)\xi}{\bar{u}(r+\xi)(r+\xi(1-\alpha))} \leq \frac{v_C}{\bar{v}} \leq \frac{\xi}{r+\xi} - \frac{\xi^2(1-\alpha)}{(r+\xi)^2}\frac{u_C}{\bar{u}}.$$

This reinforces that conflict may not be deterministic. Conditions that admit permanent conflict may also admit conceding on even difficult issues.

6 Discussion

In this paper we provide a model that predicts the dynamics of bargaining when proposal power evolves endogenously. We show that agents will compromise when the proposer is neither too strong nor too weak. A strong proposer is never willing to offer a compromise, while the responder is never willing to compromise when the proposer is too weak. The incentive constraints of both the proposer and responder determine an intermediate range of beliefs such that compromise occurs. In this interval, the proposer believes he is too weak to force an issue successfully, and the responder prefers to compromise rather than learn about the proposer's type.

We seek to understand how the evolution of power may explain puzzling bargaining outcomes. That is, observing conflict on issues that may have previously been a relatively easy issue to settle. We show that when difficult issues arise that force the agents into conflict, it also forces learning about the proposer's strength. If the proposer learns either that he is tough with certainty, and becomes too weak, then conflict ensues on every issue. But we also show the existence of an equilibrium where the proposer accommodates the responder on difficult issues when the belief about his type is too low. This avoids learning, and allows the compromise to be sustained for easy issues. We think this helps explain instances where proposers “avoid the issue”.

We considered the possibility of easy and difficult issues arising exogenously. In future work we will consider endogenous sequencing of issues according to their level of “difficulty”. Considering that learning takes place it will be useful to understand the order in which we expect to see proposals. This will likely require a model with a continuous set of possible issues. We also consider that all easy issues have the same payoff structure. It is possibile that the difficulty of the issue is related to differences in payoffs. It would also be fruitful to analyze the case of issues with heterogenous payoffs.

References

Abreu, D. and F. Gul (2000). Bargaining and Reputation. *Econometrica 68*(1), 85–117.

Admati, A. and M. Perry (1987). Strategic delay in bargaining. *The Review of Economic Studies 54*(3), 345–364.

Anesi, V. and T. Bowen (forthcoming). Policy Experimentation in Committees: a Case Against Veto Rights Under Redistributive Constraints. *American Economic Journal: Microeconomics*.

Banks, J. S. and J. Duggan (2000). A bargaining model of collective choice. *American Political Science Review 94*(1), 73–88.

Banks, J. S. and J. Duggan (2006). A general bargaining model of legislative policy-making. *Quarterly Journal of Political Science 1*(1), 49–85.

Baron, D. P. (1996). A Dynamic Theory of Collective Goods Programs. *The American Political Science Review 90*(2), 316–330.

Baron, D. P. and J. A. Ferejohn (1987). Bargaining and agenda formation in legislatures. *American Economic Review 7*(2), 303–309.

Baron, D. P. and J. A. Ferejohn (1989). Bargaining in Legislatures. *The American Political Science Review 83*(4), 1181–1206.

Bowen, T. R., Y. Chen, and H. Eraslan (2014). Mandatory Versus Discretionary Spending: The Status Quo Effect. *American Economic Review 104*(10), 2941–2974.

Bowen, T. R. and Z. Zahran (2012). On Dynamic Compromise. *Games and Economic Behavior 76*, 391–419.

Cai, H. and D. Treisman (2009). Political Decentralization and Policy Experimentation. *Quarterly Journal of Political Science 4*(1), 35–58.

Callander, S. (2011). Searching for Good Policies. *American Political Science Review 105*, 643–662.

Callander, S. and P. Hummel (2014). Preemtive Policy Experimentation. *Econometrica 82*(4), 1509–1528.

Cramton, P. (1992). Strategic Delay in Bargaining with Two-Sided Uncertainty. *The Review of Economic Studies 59*(1), 205–225.

Deneckere, R. J. and M. Y. Liang (2006). Bargaining with interdependent values. *Econometrica 74*(5), 1309–1364.

Diermeier, D. and P. Fong (2011). Legislative bargaining with reconsideration. *The Quarterly Journal of Economics 126*(2), 947–985.

Dixit, A. (2002). *The Art of Smooth Pasting*. Routledge.

Duggan, J. and T. Kalandrakis (2012). Dynamic legislative policy making. *Journal of Economic Theory 147*(5), 1653–1688.

Dziuda, W. and A. Loeper (2016). Dynamic Collective Choice with Endogenous Status Quo. *Journal of Political Economy 124*(4), 1148–1186.

Freer, M., C. Martinelli, and S. Wang (2018). Collective Experimentation: A Laboratory Study. *Working Paper*.

Friedenberg, A. (2019). Bargaining Under Strategic Uncertainty: The Role of second-Order Optimism. *Econometrica 87*(6), 1835–1865.

Fudenberg, D., D. Levine, and J. Tirole (1985). Infinite-Horizon Models of Bargaining with One-Sided Incomplete Information. In A. E. Roth (Ed.), *Game-Theoretic Models of Bargaining*, pp. 73–98. Cambridge University Press.

Hirsch, A. V. (2016). Experimentation and Persuasion in Political Organizations. *American Political Science Review 110*(1), 68–84.

Kalandrakis, T. (2004). A three-player dynamic majoritarian bargaining game. *Journal of Economic Theory 116*(2), 294–314.

Keller, G., S. Rady, and M. Cripps (2005). Strategi Experimentation with Exponential Bandits. *Econometrica 73*(1), 39–68.

Klein, N. and S. Rady (2011). Negatively Correlated Bandits. *Review of Economic Studies 78*, 693–732.

Lee, J. and Q. Liu (2013). Gambling Reputation: Repeated Bargaining with Outside Options. *Econometrica 81*(4), 1601–1672.

Majumdar, S. and S. W. Mukand (2004). Policy Gambles. *The American Economic Review 94*(4), 1207–1222.

Millner, A., H. Ollivier, and L. Simon (2014, December). Policy experimentation, political competition, and heterogeneous beliefs. *Journal of Public Economics 120*, 84–96.

Ortner, J. (2019). A continuous-time model of bilateral bargaining. *Games and Economic Behavior 113*, 720–733.

Perry, M. and P. Reny (1993). A Non-cooperatiive Bargaining Model with Strategically Timed Offers. *Journal of Economic Theory 59*, 50–77.

Powell, R. (2004). Bargaining and Learning While Fighting. *American Journal of Political Science 48*(2), 344–361.

Romer, T. and H. Rosenthal (1979). Bureaucrats versus voters: On the political economy of resource allocation by direct democracy. *The Quarterly Journal of Economics 93*(4), 563–587.

Rubinstein, A. (1982). Perfect Equilibrium in a Bargaining Model. *Econometrica 50*(1), 97–109.

Strulovici, B. (2010). Learning While Voting: Determinants of Collective Experimentation. *Econometrica 78*(3), 933–971.

Volden, C., M. M. Ting, and D. P. Carpenter (2008). A Formal Model of Learning and Policy Diffusion. *The American Political Science Review 102*(3), 319–332.

7 Appendix

7.1 Proof of Proposition 2

The proof is conducted as follows. First, we derive the value function of each player under the conjectured strategy profile. Then we verify each player's incentive conditions and derive the values of $\underline{p}$ and $\overline{p}$.

Value Functions Observe that $V_i(p) = V_{i,c}(p)$ for any $p < \underline{p}$, since the belief dynamics imply that if $p_{t'}\underline{p}$ for some $t' > 0$, the players will always create conflict for any $t > t'$. Similarly, $V_i(p) = V_{i,n}(p)$ for any $p \in [\underline{p}, \overline{p}]$. Therefore, it remains to derive the value functions for $p > \overline{p}$.

First, consider the proposer's value function $V_P(p)$ for $p > \overline{p}$. In the interval $[t, t + dt)$, the probability that the proposer overturns a rejection is given by $1 - e^{-\lambda dt} = \lambda dt$. In this case, the issue is resolved, and the proposer receives utility $\bar{u}$. In addition, the proposer's type is now known to be high, i.e., $p = 1$. A new issues arrives in the interval $[t, t + dt)$ with probability $1 - e^{-\xi dt} = \xi dt$. In this case, we switch to the value function $V_P(1)$. With the remaining probability, $1 - \xi dt$, the players wait for a new issue to arrive, and hence the value function becomes $\hat{V}_P(1)$. With probability $1 - p\lambda dt$ the issue is not resolved in $[t, t + dt)$. As a consequence, the belief decreases to $p + dp$, where dp will be computed below (and of course $dp < 0$). The continuation utility in this case is $V_P(p + dp)$.

Therefore, $V_P(p)$ is determined recursively by the equation

$$V_P(p) = p\lambda dt \left(\bar{u} + e^{-rdt}(\xi dt V_P(1) + (1 - \xi dt)\hat{V}_P(1))\right) + (1 - p\lambda dt)e^{-rdt} V_P(p + dp),$$

where

$$\hat{V}_i(p) = \frac{\xi}{r + \xi} V_i(p).$$

Substituting $1 - rdt$ for e^{-rdt} yields

$$V_P(p) = p\lambda dt \left(\bar{u} + (1 - rdt)(\xi dt V_P(1) + (1 - \xi dt)\hat{V}_P(1))\right) + (1 - p\lambda dt)(1 - rdt)V_P(p + dp).$$

Note that $V_P(p + dp) = V_P(p) + V_P'(p)dp$, and the belief dynamics are given by $dp/dt = -\lambda p(1 - p)$. Therefore,

$$\begin{aligned}(1 - (1 - p\lambda dt)(1 - rdt))V_P(p) = &\ p\lambda dt \left(\bar{u} + (1 - rdt)(\xi dt V_P(1) + (1 - \xi dt)\hat{V}_P(1))\right) \\ &- (1 - p\lambda dt)(1 - rdt)V_P'(p)\lambda p(1 - p)dt.\end{aligned}$$

Dropping higher order terms, i.e., dt^2 and dt^3 we get

$$(r + p\lambda)V_P(p)dt = p\lambda \left(\bar{u} + \hat{V}_P(1)\right) dt - V_P'(p)\lambda p(1 - p)dt.$$

Substituting the value of $\hat{V}_P(1)$ yields the differential equation

$$(r + p\lambda)V_P(p) = p\lambda\bar{u} \left(1 + \frac{\xi\lambda}{r(r + \lambda + \xi)}\right) - V_P'(p)\lambda p(1 - p).$$

Solving the differential equation we get

$$V_P(p) = \frac{(1-p)^{1+\frac{r}{\lambda}}}{p^{\frac{r}{\lambda}}} K_P + V_{P,c}(p), \tag{13}$$

where K_P is the integration constant, and $V_{P,c}(p)$ is the permanent-conflict value function derived in Section 4. We can see that $V_P(p)$ is striclty convex. In particular, $V_{P,c}(p)$ is affine linear, $K_P > 0$ and the first summand in (13) is therefore strictly convex.

The boundary condition is $V_P(\overline{p}) = V_{P,n}$, where $V_{p,n}$ is the continuation value if there is no conflict, derived in section 7.4. It is easy of verify that $V_{P,n} > V_{P,c}(\overline{p})$ if and only if $u_C/\bar{u} < \frac{\lambda}{r+\lambda+\xi}$. Hence $K > 0$.

We now determine the responder's value function $V_R(p)$ for $p > \overline{p}$. Using a similar argument as for the proposer, the responder's value function is defined recursively as follows:

$$V_R(p) = p\lambda dt e^{-rdt}(\xi dt V_R(1) + (1-\xi dt)\hat{V}_R(1)) + (1-p\lambda dt)\left(\xi dt\bar{v} + e^{-rdt}V_R(p+dp)\right).$$

Applying $1 - rdt = e^{-rdt}$ and $V_R(p+dp) = V_R(p) + V_R'(p)dp$ yield

$$\begin{aligned}(1-(1-p\lambda dt)(1-rdt))\, V_R(p) =& p\lambda dt(1-rdt)(\xi dt V_R(1) + (1-\xi dt)\hat{V}_R(1)) \\ &+ (1-p\lambda dt)\left(\xi dt\bar{v} - (1-rdt)V_R'(p)\lambda p(1-p)dt\right).\end{aligned}$$

Dropping higher order terms, we have

$$(r + p\lambda)V_R(p)dt = p\lambda\hat{V}_R(1)dt + \xi\bar{v}dt - V_R'(p)\lambda p(1-p)dt.$$

Thus, V_R satisfies the following differential equation

$$(r + p\lambda)V_R(p) = \frac{p\lambda\xi^2\bar{v}}{r(r+\lambda+\xi)} + \xi\bar{v} - V_R'(p)\lambda p(1-p).$$

Solving the differential equation yields

$$V_R(p) = \frac{(1-p)^{1+\frac{r}{\lambda}}}{p^{\frac{r}{\lambda}}} K_R + V_{R,c}(p), \tag{14}$$

where K_R is the integration constant. The boundary condition that determines K_R is given by $V_R(\overline{p}) = V_{R,n}$. Below, we show that K_R is positive.

Equilibrium verification Given the above value functions, we verify that the equilibrium profile is indeed optimal for each player.

For $p < \underline{p}$, the responder rejects any offer from the proposer, and the proposer offers x_P. Given the responder's behavior, the proposer's incentive condition is trivial. Therefore, it suffices to check if the responder rejects a compromise offer x_c if the proposer deviates and makes such proposal. This requires that

$$V_{R,c}(p) \geq v_C + \hat{V}_{R,c}(p)$$

Note that this inequality simplifies to $V_{R,c}(p) \geq V_{R,n}$. Since $V_{R,c}(p)$ is decreasing in p, the above inequality satisfies if and only if

$$V_{R,c}(\underline{p}) \geq V_{R,n}. \tag{15}$$

Next, we check each player's incentive for $p \in \bar{C}$. First consider the responder's incentives: We need to check two incentive conditions: (a) responder's incentive to accept x_c; and (b) his incentive to reject x_P.

If the responder accepts the the proposer's offer x_c, then the belief stays the same and there will be an agreement at x_c for all future periods. If the responder rejects the offer, then learning occurs. With probability $p\lambda dt$ the proposer overturns the rejection and the issue is resolved with position x_P. With the complementary probability, however, the proposer fails to overturn, in which case the current is continues and the belief gradually declines. Therefore, the responder is better off accepting the offer x_c if

$$V_R(p) \geq p\lambda dt e^{-rdt}(\xi dt V_R(1) + (1-\xi dt)\hat{V}_R(1)) + (1-p\lambda dt)\left(\xi dt\bar{v} + e^{-rdt}V_R(p+dp)\right). \tag{16}$$

There are two cases to consider:

Case 1: The belief after the proposer's failure moves outside the compromise region, i.e., $p + dp < \underline{p}$. In this case, the incentive condition (16) becomes

$$V_{R,n} \geq p\lambda dt e^{-rdt}(\xi dt V_R(1) + (1-\xi dt)\hat{V}_R(1)) + (1-p\lambda dt)\left(\xi dt\bar{v} + e^{-rdt}V_{R,c}(\underline{p})\right).$$

Note the last term: If the responder rejects x_P and the proposer fails to overturn, then the belief goes out of the compromise region. Dropping the terms with orders of dt and higher, we have

$$V_{R,n} \geq V_{R,c}(\underline{p}). \tag{17}$$

Combining with (15), we have $V_{R,n} = V_{R,c}(\underline{p})$, which gives the formula for the lower bound:

$$\underline{p} = \frac{(\xi\bar{v} - (\xi + r)v_C)(r + \lambda + \xi)}{\xi\bar{v}\lambda}. \tag{18}$$

Case 2: The belief after the proposer's failure is still in the compromise region, i.e., $p+dp \geq \underline{p}$. In this case, (16) becomes

$$V_{R,n} \geq p\lambda dt e^{-rdt}(\xi dt V_R(1) + (1-\xi dt)\hat{V}_R(1)) + (1-p\lambda dt)\left(\xi dt\bar{v} + e^{-rdt}V_{R,n}\right). \tag{19}$$

Dropping terms with orders of dt^2 and higher, we get

$$p\lambda\left(V_{R,n} - \hat{V}_R(1)\right) + rV_{R,n} \geq \xi\bar{v} \tag{20}$$

Note that the left-hand side (the right-hand side) of the above inequality captures the cost (benefit) of rejecting x_c. First, the cost: when the proposer succeeds in overturning the rejection, then the future value function drops to $\hat{V}_R(1)$. Further, in case of failure, the agreement is delayed:

this is captured with the term $rV_{R,n}$. The benefit comes from the case in which the Proposer fails to overturn the rejection and the current is replaced with the new one.

It remains to show that (20) holds for all $p \in [\underline{p}, \overline{p}]$. Given that the right-hand side of (20) is decreasing in p, it is sufficient to show that the inequality holds at $p = \underline{p}$. Plugging in (18), it can be shown that (20) holds at $p = \underline{p}$ if and only if

$$\frac{v_C}{\bar{v}} > \frac{\xi}{r + \lambda + \xi}.$$

Since $\underline{p} < 1$ if and only if $\frac{v_C}{\bar{v}} > \frac{\xi}{r+\lambda+\xi}$, it must be that (20) holds for all $p \in [\underline{p}, \overline{p}]$ as long as $\underline{p} < 1$.

Next, the responder must reject x_P when it is offered by the proposer. Again, we have the following two cases:

Case 1: The belief after the proposer's failure moves outside the compromise region. In this case, the responder's incentive condition holds if $V_{R,c}(\underline{p}) \geq \hat{V}_{R,n}$. The continuation payoff from rejecting the offer is $V_{R,c}(\underline{p})$ because the belief moves outside the compromise set. If, instead, the responder accepts the offer, then the belief remains in the compromise set and the continuation payoff is $\hat{V}_{R,n}$. This inequality trivially holds, because $V_{R,c}(\underline{p}) = V_{R,n} \geq \hat{V}_{R,n}$.

Case 2: The belief after the proposer's failure is still in the compromise region. In this case, the responder's incentive constraint holds if $V_{R,n} \geq \hat{V}_{R,n}$, which is always satisfied.

Next, consider the proposer's incentive constraints for $p \in C$. In region C the proposer should prefer offering x_c (which is immediately accepted) to offering x_P (which is rejected). If the belief after a proposer's failure remains in the compromise set, then

$$V_{P,n} \geq p\lambda dt\left(\bar{u} + e^{-rdt}(\xi dt V_P(1) + (1 - \xi dt)\hat{V}_P(1)\right) + (1 - p\lambda dt)e^{-rdt}V_{P,n}. \tag{21}$$

Solving (21) yields

$$p\lambda\left(\bar{u} + (\hat{V}_P(1) - V_{P,n})\right) \leq rV_{P,n}. \tag{22}$$

Similar to the responder's incentive constraint, both terms in the inequality capture the proposer's benefit and cost of creating conflict (by offering x_c). The benefit comes from overturning the rejection, in which case the proposer enjoys not only the lump-sum payoff of $\bar{u}$ but also an increase in the value function. However, the cost comes in the case of failure, because the agreement is delayed.

(22) determines the upper bound, $\overline{p}$, of the compromise set. In order for $\overline{p} < 1$, the inequality must be violated for $p = 1$ and hold for smaller p. In order for (22) to be violated at $p = 1$ we must have $\bar{u} + \hat{V}_P(1) > \frac{r+\lambda}{\lambda}V_{P,n}$, which simplifies to

$$\frac{u_C}{\bar{u}} < \frac{\lambda}{r + \lambda + \xi}. \tag{23}$$

For $p = 0$ the equation always holds. Therefore, given that (23) holds, (22) implies that the upper bound of the compromise set is

$$\overline{p} = \frac{rV_{P,n}}{\lambda\left(\bar{u} + \hat{V}_P(1) - V_{P,n}\right)}, \tag{24}$$

which is equivalent to

$$\overline{p} = \frac{r(r+\lambda+\xi)u_C}{\lambda((r+\lambda)\bar{u} - (r+\lambda+\xi)u_C)}. \tag{25}$$

In order for the compromise set to be non-empty it must be the case that $\underline{p} < \overline{p}$.

Finally, suppose by way of contradiction that it is optimal for the proposer to deviate and offer x_c when $p > \overline{p}$. This deviation is only optimal if the responder would accept x_c. However, then it should be optimal to for the proposer to offer x_c in every period, which is not the case when $p > \overline{p}$, a contradiction.

7.2 Proof of Proposition 3

A simple calculation yields

$$\underline{p} - \underline{p}_1 = \frac{[\bar{v}\xi - v_C(r+\xi)][v_C(r+\lambda+\xi) - \bar{v}\xi]}{\xi\lambda\bar{v}v_C},$$

which is strictly positive if

$$\frac{\xi}{r+\lambda+\xi} < \frac{v_C}{\overline{v}} < \frac{\xi}{r+\xi}. \tag{26}$$

Also,

$$\overline{p}_1 - \overline{p} = \frac{u_C\xi[\overline{u}\lambda - u_C(\lambda+r+\xi)]}{\lambda(\overline{u}-u)[\overline{u}(\lambda+r) - u_C(\lambda+r+\xi)]},$$

which is strictly positive if

$$\frac{u_C}{\overline{u}} < \frac{\lambda}{\lambda+r+\xi}. \tag{27}$$

Conditions (26) and (27) are implied by the conditions in Corollary 2.4. This proves the result.

7.3 Proof of Propositions 5 and 6

We first derive the players' value functions in each type of equilibrium. Observe that the two types of equilibrium differs only in the behavior at $p = \underline{p}_{\alpha}$. Therefore, the differential equations underlying both value functions are identical, and they differ only in the boundary conditions at $p = \underline{p}_{\alpha}$. After obtaining the value functions, we verify each type of equilibrium by investigating the players' incentive conditions.

Value function Let $V_i(p)$, $i = P, C$ be the value function under the easy issues, and let $W_i(p)$ be the value function under the difficult issues. Also, for notational simplicity, define $Z_i(p)$ to be the value function when the new issue arises, and define $\hat{Z}_i(p)$ as the value function when the current issue is resolved (but the new issue has not yet appeared). Then it is straightforward that

$$Z_i(p) = (1-\alpha)V_i(p) + \alpha W_i(p), \tag{28}$$

$$\hat{Z}_i(p) = \frac{\xi}{r+\xi}\Big((1-\alpha)V_i(p) + \alpha W_i(p)\Big). \tag{29}$$

Observe that for $p < \underline{p}_\alpha$, regardless of the policy issue type, a permanent conflict occurs in equilibrium. Therefore, $V_i(p) = W_i(p) = V_{i,c}(p)$.

Next, we derive the value functions for $p \in (\underline{p}_\alpha, \overline{p}_\alpha]$. Consider first the proposer's value function. Under the easy issue, the proposer offers x_c and the responder accepts the offer. Therefore, $V_P(p)$ satisfies

$$V_P(p) = u_C + e^{-rdt}\Big(\xi dt Z_P(p) + (1-\xi dt)\hat{Z}_P(p)\Big).$$

Cancelling the terms with order dt or higher and applying (28) and (29), we have

$$V_P(p) = u_C + \frac{\xi}{r+\xi}\Big((1-\alpha)V_P(p) + \alpha W_P(p)\Big). \tag{30}$$

Under the difficult issue, the proposer offers x_P and the responder rejects the offer. Therefore, $W_P(p)$ satisfies

$$\begin{aligned} W_P(p) =& p\lambda dt\Big(\bar{u} + e^{-rdt}(\xi dt V_{P,c}(1) + (1-\xi dt)\hat{V}_{P,c}(1))\Big) \\ &+ (1-p\lambda dt)e^{-rdt}\Big(\xi dt Z_P(p+dp) + (1-\xi dt)W_P(p+dp)\Big). \end{aligned} \tag{31}$$

Cancelling the terms with order dt^2 or higher, applying (28) and (29), and reorganizing yields

$$\lambda p(1-p)W_P'(p) = p\lambda(\bar{u} + \hat{V}_{P,c}(1)) + \xi(1-\alpha)V_P(p) - (p\lambda + r + \xi(1-\alpha))W_P(p). \tag{32}$$

Equations (30) and (32) jointly determine the proposer's value functions.

Now consider the responder. Similar to (30), the responder's value function under the easy issues are given by

$$V_R(p) = v_C + \frac{\xi}{r+\xi}\Big((1-\alpha)V_R(p) + \alpha W_R(p)\Big). \tag{33}$$

Under the difficult issue, the proposer offers x_P and the responder rejects the offer. Therefore, $W_R(p)$ satisfies

$$\begin{aligned} W_R(p) =& p\lambda dt e^{-rdt}(\xi dt V_{R,c}(1) + (1-\xi dt)\hat{V}_{R,c}(1)) \\ &+ (1-p\lambda dt)\Big(\xi dt\bar{v} + e^{-rdt}(\xi dt Z_R(p+dp) + (1-\xi dt)W_R(p+dp))\Big). \end{aligned}$$

Cancelling the terms with order dt^2 or higher, applying (28) and (29), and reorganizing yields

$$\lambda p(1-p)W_R'(p) = p\lambda\hat{V}_{R,c}(1) + \xi\bar{v} + \xi(1-\alpha)V_R(p) - (p\lambda + r + \xi(1-\alpha))W_R(p). \tag{34}$$

Equations (33) and (34) jointly determine the proposer's value functions.

For $p > \overline{p}_\alpha$, conflict arises in both the easy and the difficult policy issues. For the proposer, $V_P(p)$ satisfies

$$\begin{aligned} V_P(p) =& p\lambda dt\Big(\bar{u} + e^{-rdt}(\xi dt V_{P,c}(1) + (1-\xi dt)\hat{V}_{P,c}(1))\Big) \\ &+ (1 - p\lambda dt)e^{-rdt}\Big(\xi dt Z_P(p+dp) + (1-\xi dt)V_P(p+dp)\Big), \end{aligned}$$

and $W_P(p)$ satisfies (31). Simplifying, we have

$$\begin{aligned} \lambda p(1-p)V_P'(p) &= p\lambda(\bar{u} + \hat{V}_{P,c}(1)) + \xi\alpha W_P(p) - (p\lambda + r + \xi\alpha)V_P(p). \\ \lambda p(1-p)W_P'(p) &= p\lambda(\bar{u} + \hat{V}_{P,c}(1)) + \xi(1-\alpha)V_P(p) - (p\lambda + r + \xi(1-\alpha))W_P(p). \end{aligned}$$

Similarly, $V_R(p)$ and $W_R(p)$ jointly solve

$$\begin{aligned} \lambda p(1-p)V_R'(p) &= p\lambda\hat{V}_{R,c}(1) + \xi\bar{v} + \xi\alpha V_R(p) - (p\lambda + r + \xi\alpha)W_R(p). \\ \lambda p(1-p)W_R'(p) &= p\lambda\hat{V}_{R,c}(1) + \xi\bar{v} + \xi(1-\alpha)V_R(p) - (p\lambda + r + \xi(1-\alpha))W_R(p). \end{aligned}$$

Equilibrium verification: Conflict equilibrium In the first type of equilibrium—*conflict equilibrium*—the proposer induces conflict for any p when the current policy issue is a difficult one. Therefore, if players currently faces a difficult policy issue and $p = \underline{p}_\alpha$, there will be a permanent conflict. Thus, the boundary conditions for W_i are given by $W_i(\underline{p}_\alpha) = V_{i,c}(\underline{p}_\alpha)$. Then from (30) and (33), we have

$$V_P(\underline{p}_\alpha) = \frac{(r+\xi)u_C + \xi\alpha V_{P,c}(\underline{p}_\alpha)}{r+\xi\alpha} \tag{35}$$

$$V_R(\underline{p}_\alpha) = \frac{(r+\xi)v_C + \xi\alpha V_{R,c}(\underline{p}_\alpha)}{r+\xi\alpha} \tag{36}$$

To fix the value of $\underline{p}_\alpha$, first consider the responder's incentives. First, under easy policy issues, the responder must accept x_c at $p = \underline{p}_\alpha$. This condition is given by

$$\begin{aligned} &v_C + e^{-rdt}(\xi dt Z_R(\underline{p}_\alpha) + (1-\xi dt)\hat{Z}_R(\underline{p}_\alpha)) \\ &\quad \geq p\lambda dt e^{-rdt}(\xi dt V_{R,c}(1) + (1-\xi dt)\hat{V}_{R,c}(1)) + (1-p\lambda dt)\Big(\xi dt\bar{v} + e^{-rdt}V_{R,c}(\underline{p}_\alpha + dp)\Big) \end{aligned}$$

Deleting terms with order dt or higher and reorganizing yield

$$v_C + \frac{\xi}{r+\xi}((1-\alpha)V_R(\underline{p}_\alpha) + \alpha W_R(\underline{p}_\alpha)) \geq V_{R,c}(\underline{p}_\alpha).$$

Plugging in $W_R(\underline{p}_\alpha) = V_{R,c}(\underline{p}_\alpha)$ and (36), and simplifying yield

$$V_{R,n} \geq V_{R,c}(\underline{p}_\alpha).$$

Interestingly, the incentive condition at the lower bound is identical to the one in the benchmark model.

For $p < \underline{p}_\alpha$, the responder's incentive condition to reject x_c is the same as the benchmark model, which is

$$V_{R,n} \le V_{R,c}(p) \qquad \text{for any } p < \underline{p}_\alpha$$

Combining the above two incentive conditions yields that

$$V_{R,n} = V_{R,c}(\underline{p}_\alpha),$$

which is identical to the benchmark model. Therefore,

$$\underline{p}_\alpha = \underline{p} = \frac{(\xi\bar{v} - (\xi + r)v_C)(r + \lambda + \xi)}{\xi\bar{v}\lambda}.$$

Next, consider the proposer's incentive to offer x_P at $p = \underline{p}_\alpha$. Because he must prefer offering x_P to offering x_0, his incentive condition is given by

$$V_{P,c}(\underline{p}_\alpha) \ge e^{-rdt}(\xi dt Z_P(\underline{p}_\alpha) + (1 - \xi dt)\hat{Z}_P(\underline{p}_\alpha)).$$

Reorganizing yields

$$V_{P,c}(\underline{p}_\alpha) \ge \frac{\xi}{r + \xi}((1 - \alpha)V_P(\underline{p}_\alpha) + \alpha V_{P,c}(\underline{p}_\alpha))$$

Plugging in (35) and simplifying again, we have

$$V_{P,c}(\underline{p}_\alpha) \ge \frac{\xi(1 - \alpha)}{r}u_C.$$

Plugging in the value of $\underline{p}_\alpha$ and simplifying yields the condition for the equilibrium existence:

$$\frac{\xi(1 - \alpha)}{r + \xi}\frac{u_C}{\bar{u}} + \frac{r + \xi}{\xi}\frac{v_C}{\bar{v}} \le 1.$$

Equilibrium verification: "Avoiding the issue" equilibrium In the second type of equilibrium—*"avoiding the issue" equilibrium*—the proposer offers x_0 to induce compromise at $p = \underline{p}_\alpha$ under difficult policy issues. Then, similar to the argument above, the value of $V_P(\underline{p}_\alpha)$ and $W_P(\underline{p}_\alpha)$ satisfy the following system of equations:

$$V_P(\underline{p}_\alpha) = u_C + \frac{\xi}{r + \xi}\Big((1 - \alpha)V_P(\underline{p}_\alpha) + \alpha W_P(\underline{p}_\alpha)\Big)$$
$$W_P(\underline{p}_\alpha) = \frac{\xi}{r + \xi}\Big((1 - \alpha)V_P(\underline{p}_\alpha) + \alpha W_P(\underline{p}_\alpha)\Big)$$

Similarly, $V_R(\underline{p}_\alpha)$ and $W_R(\underline{p}_\alpha)$ jointly solve

$$V_R(\underline{p}_\alpha) = v_C + \frac{\xi}{r + \xi}\Big((1 - \alpha)V_R(\underline{p}_\alpha) + \alpha W_R(\underline{p}_\alpha)\Big)$$
$$W_R(\underline{p}_\alpha) = \bar{v} + \frac{\xi}{r + \xi}\Big((1 - \alpha)V_R(\underline{p}_\alpha) + \alpha W_R(\underline{p}_\alpha)\Big)$$

Solving the above systems yields the boundary conditions at $p = \underline{p}_\alpha$:

$$V_P(\underline{p}_\alpha) = u_C + \frac{\xi(1-\alpha)}{r}u_C, \qquad V_R(\underline{p}_\alpha) = v_C + \frac{\xi}{r}((1-\alpha)v_C + \alpha\bar{v}),$$
$$W_P(\underline{p}_\alpha) = \frac{\xi(1-\alpha)}{r}u_C, \qquad W_R(\underline{p}_\alpha) = \bar{v} + \frac{\xi}{r}((1-\alpha)v_C + \alpha\bar{v}).$$

Having fixed the value functions, let us consider the responder's incentives. First, under easy policy issues, the responder must accept x_c at $p = \underline{p}_\alpha$. Same as the first type of equilibrium, this condition is given by

$$\begin{aligned} &v_C + e^{-rdt}(\xi dt Z_R(\underline{p}_\alpha) + (1-\xi dt)\hat{Z}_R(\underline{p}_\alpha)) \\ &\quad \geq p\lambda dt e^{-rdt}(\xi dt V_{R,c}(1) + (1-\xi dt)\hat{V}_{R,c}(1)) + (1-p\lambda dt)\Big(\xi dt\bar{v} + e^{-rdt}V_{R,c}(\underline{p}_\alpha + dp)\Big) \end{aligned}$$

Deleting terms with order dt or higher and reorganizing yield

$$v_C + \frac{\xi}{r+\xi}((1-\alpha)V_R(\underline{p}_\alpha) + \alpha W_R(\underline{p}_\alpha)) \geq V_{R,c}(\underline{p}_\alpha).$$

Plugging in the value of $W_R(\underline{p}_\alpha)$ and $V_R(\underline{p}_\alpha)$ and simplifying yield

$$V_{R,n} + \frac{\xi\alpha}{r}(\bar{v} - v_C) \geq V_{R,c}(\underline{p}_\alpha).$$

Note that the left-hand side of the incentive condition is greater than that in the first-type of equilibrium.

For $p < \underline{p}_\alpha$, the incentive condition to reject x_c is identical to the benchmark model (and the first type of equilibrium), which is

$$V_{R,n} \leq V_{R,c}(p) \qquad \text{for any } p < \underline{p}_\alpha$$

Combining the above two incentive conditions yields show that the responder's incentive condition is satisfied for any $\underline{p}_\alpha \in [\underline{p}_0, \underline{p}]$, where

$$\underline{p}_0 = \frac{(r+\lambda+\xi)(\xi(1-\alpha)\bar{v} - (r+\xi(1-\alpha))v_C)}{\lambda\xi\bar{v}},$$

and

$$\underline{p} = \frac{(r+\lambda+\xi)(\xi\bar{v} - (r+\xi)v_C)}{\lambda\xi\bar{v}},$$

which is identical of the lower bound in the benchmark model.

Next, consider the proposer's incentive to offer x_0 at $p = \underline{p}_\alpha$. Since he must prefer offering x_0 than offering x_P, his incentive condition is given by

$$V_{P,c}(\underline{p}_\alpha) \leq e^{-rdt}(\xi dt Z_P(\underline{p}_\alpha) + (1-\xi dt)\hat{Z}_P(\underline{p}_\alpha)).$$

Reorganizing yields

$$V_{P,c}(\underline{p}_\alpha) \le \frac{\xi}{r+\xi}((1-\alpha)V_P(\underline{p}_\alpha) + \alpha W_P(\underline{p}_\alpha))$$

Plugging in the value of $W_P(\underline{p}_\alpha)$ and $V_P(\underline{p}_\alpha)$ and simplifying again, we have

$$V_{P,c}(\underline{p}_\alpha) \le \frac{\xi(1-\alpha)}{r} u_C,$$

which is identical to the incentive condition in the first type of equilibrium, except the direction of the inequality reversed. As $\underline{p}_\alpha$ decreases, the above inequality becomes weaker (i.e., it satisfies under a broader range of parameter). Therefore, the incentive condition is the weakest if we set $\underline{p}_\alpha = \underline{p}_0$. Plugging in $\underline{p}_\alpha = \underline{p}_0$ and simplifying yield

$$\frac{\xi(1-\alpha)}{r+\xi}\frac{u_C}{\bar{u}} + \frac{r+\xi(1-\alpha)}{\xi}\frac{v_C}{\bar{v}} \ge 1-\alpha.$$

7.4 Model without delay

A Proposer (P) and Responder (R) repeatedly bargain over multiple issues. Time $t \in [0,\infty)$ is continuous and the horizon is infinite. The arrival rate of a new issue is ξ. Thus over time interval $[t, t+dt)$ a new issue arrives with probability $1 - e^{-\xi dt} \approx \xi dt$. When a new issue arrives, the Proposer announces a proposal $x \in X = \{x_0, x_C, x_P\}$. We interpret x_0 as the status-quo position, x_C as the compromise position, and x_P as the Proposer's most preferred position. After the proposal is made, the responder chooses whether to accept or reject x. If the responder accepts, then the issue is resolved with position x. If the responder rejects, then the proposer chooses whether to attempt to overturn the rejection of x or not. If P does not attempt to overturn, then the issue is resolved with the status quo position x_0. If P attempts to overturn then the issue is overturned at rate $\theta\lambda$. If the issue is overturned, then it is resolved with position x. If P chooses to overturn and is not successful before the next issue arrives, then the issue is not resolved and position x_0 remains. We assume all decisions happen instantaneously and all payoffs are realized when the new issue arrives.

Whether or not the rejection is overturned depends on the proposer's ability $\theta \in \{0,1\}$, which can be either high $\theta = 1$ or low $\theta = 0$. In particular, if the proposer's ability is high, then R's rejection is overturned at rate λ. Otherwise, if his ability is low then a rejection is never overturned. Both players are symmetrically uninformed about the proposer's ability. Let p_t be the common prior belief at time t that the proposer's ability is high. The assumption that the proposer has the same information about his ability as the responder applies if all parties are symmetrically informed about the proposer's previous bargaining skills.

For simplicity we assume that players receive utility from the issue at the time when the current issue is replaced by the new one. The utility depends on what position the issue was resolved in. If the issue has been resolved with position x, then the payoffs are $u_P(x)$ and $u_R(x)$. If the issue has not been resolved, then the status quo is retained and utilities are $u_P(x_0)$ and $u_R(x_0)$.

We assume that R strictly prefers x_0 to x_P, and similarly, P strictly prefers x_P to x_0. We normalize utility of the players' such that $u_P(x_0) = u_R(x_P) = 0$. Also, let $u_P(x_P) = \bar{u}$, $u_P(x_C) = u_C$, $u_R(x_0) = \bar{v}$ and $u_R(x_C) = v_C$. We assume that $0 < u_C, v_C$. Finally, utility is discounted at a rate $r > 0$.

We show that without delay there exist a qualitatively similar equilibrium to the case of delay. The key difference is the calculation of the bounds on p.

Proposition 7 *There exists an equilibrium when there are multiple issues in the model with no delay as follows.*

$$\chi(p) = \begin{cases} x_c & \text{if } p \in [\underline{p}, \overline{p}] \\ x_P & \text{otherwise.} \end{cases} \qquad A(p) = \begin{cases} \{x_0, x_c\} & \text{if } p \geq \underline{p} \\ \{x_0\} & \text{if } p < \underline{p}, \end{cases}$$

where

$$\begin{aligned} \underline{p} &= \frac{(\bar{v} - v_C)(r + \lambda + \xi)}{\bar{v}\lambda} \\ \overline{p} &= \frac{u_C r(\lambda + r + \xi)}{[(\bar{u} - u_C)(r + \lambda) - \xi u_C]\lambda}. \end{aligned}$$

The proof follows exactly as before, but with slightly modified payoffs.

Permanent Conflict We first derive the recursive payoffs that would arise if $p = 1$ and there is permanent conflict, i.e., the proposer offers x_P in every period, which the responder always rejects. The issue only becomes resolved if the proposer is able to overturn the rejection, before a new issue arrives. This benchmark case turns out to be an important building block of the equilibrium value function.

Define $V_{i,success}(1)$ to be the player i's value function if the issue has successfully been overturned and a new issue is being awaited when $p = 1$. The next issue arrives which occurs with probability ξdt on the time interval $[t, t + dt)$. Then

$$\begin{aligned} V_{P,\text{success}}(1) &= \xi dt[\bar{u} + e^{-rdt} V_{P,\text{challenge}}(1)] + (1 - \xi dt)e^{-rdt} V_{P,\text{success}}(1) \\ V_{R,\text{success}}(1) &= e^{-rdt}[\xi dt V_{R,\text{challenge}}(1) + (1 - \xi dt) V_{R,\text{success}}(1)] \end{aligned}$$

Using the fact that $e^{-rdt} \approx 1 - rdt$ and dropping higher order terms implies

$$\begin{aligned} V_{P,\text{success}}(1) &= \frac{\xi}{\xi + r}\left(\bar{u} + V_{P,\text{challenge}}(1)\right), \\ V_{R,\text{success}}(1) &= \frac{\xi}{\xi + r} V_{R,\text{challenge}}(1). \end{aligned}$$

If the issue has not successfully been overturned, but has been rejected and being challenged, then with probabiilty ξdt the issue is replaced with the new one, and with probability λdt the challenge is successful. We denote these value functions by $V_{P,challenge}(1)$ and $V_{R,challenge}(1)$, respectively.

$$\begin{aligned} V_{P,\text{challenge}}(1) &= \lambda dt e^{-rdt} V_{P,\text{success}}(1) + (1-\lambda dt)e^{-rdt} V_{P,\text{challenge}}(1) \\ V_{R,\text{challenge}}(1) &= \lambda dt e^{-rdt} V_{R,\text{success}}(1) + \xi dt[\bar{v} + e^{-rdt} V_{R,\text{challenge}}(1)] + (1-\lambda dt - \xi dt)e^{-rdt} V_{R,\text{challenge}}(1) \end{aligned}$$

These together imply:

$$V_{P,\text{success}}(1) = \frac{\xi(\lambda + r)}{r(r+\lambda+\xi)}\bar{u}, \tag{37}$$

$$V_{R,\text{success}}(1) = \frac{\xi^2}{r(r+\lambda+\xi)}\bar{v}, \tag{38}$$

$$V_{P,\text{challenge}}(1) = \frac{\xi\lambda}{r(r+\lambda+\xi)}\bar{u}, \tag{39}$$

$$V_{R,\text{challenge}}(1) = \frac{\xi(r+\xi)}{r(r+\lambda+\xi)}\bar{v}. \tag{40}$$

When $p = 0$, the proposer never overturns the rejection (that is, $\lambda = 0$ in equations (39) and (40)), and thus the value functions are

$$V_{P,challenge}(0) = 0, \tag{41}$$

$$V_{R,challenge}(0) = \frac{\xi}{r}\bar{v}. \tag{42}$$

The player i's value function $V_{i,challenge}(p)$ is the convex combination of $V_{i,challenge}(1)$ and $V_{i,challenge}(0)$. To understand this, note that the players' never change their actions in the future: Regardless of the outcome, P always proposes x_P which R rejects. Therefore,

$$V_{P,challenge}(p) = pV_{P,c}(1) + (1-p)V_{P,c}(0) = p\bar{u}\frac{\lambda\xi}{r(r+\lambda+\xi)}, \tag{43}$$

$$V_{R,challenge}(p) = pV_{R,c}(1) + (1-p)V_{R,c}(0) = \frac{\xi\bar{v}}{r}\left(1 - \frac{p\lambda}{r+\lambda+\xi}\right). \tag{44}$$

No Conflict The second important benchmark case is where no conflict occurs. In this case, the proposer offers the compromise outcome x_C, and the responder accepts the offer x_0 and x_C but rejects x_P.

No learning occurs, and hence the payoffs do not depend on the belief p.

$$\begin{aligned} V_{P,\text{compromise}} &= \xi dt[u_C + e^{-rdt} V_{P,\text{compromise}}] + (1-\xi dt)e^{-rdt} V_{P,\text{compromise}} \\ V_{R,\text{compromise}} &= \xi dt[v_C + e^{-rdt} V_{R,\text{compromise}}] + (1-\xi dt)e^{-rdt} V_{R,\text{compromise}}] \end{aligned}$$

Simplifying gives:

$$V_{P,\text{compromise}} = \frac{\xi}{r}u_C, \tag{45}$$

$$V_{R,\text{compromise}} = \frac{\xi}{r}v_C. \tag{46}$$

Incentive compatibility follows exactly as before. Below we characterize the key bounds.

Upper bound on p We can find $\bar{p}$ as it is the p such that the proposer is just indifferent between conflict and compromise. This is

$$\frac{\xi}{r}u_C = \lambda p dt e^{-rdt} V_{P,success} + (1 - \lambda p dt)e^{-rdt}\frac{\xi}{r}u_C.$$

The above equation is given by the smooth pasting condition. Simplifying gives

$$\bar{p} = \frac{u_C r(\lambda + r + \xi)}{[(\bar{u} - u_C)(r + \lambda) - \xi u_C]\lambda}.$$

This is less than one as long as

$$\frac{u_C}{\bar{u}} \leq \frac{\lambda}{\lambda + r + \xi}. \tag{47}$$

Lower bound on p The lower bound on p is such that the responder is just indifferent between compromise and conflict and this gives

$$\frac{\xi}{r}v_C = V_{R,challenge}(p),$$

and hence,

$$\underline{p} = \frac{(\bar{v} - v_C)(r + \lambda + \xi)}{\bar{v}\lambda}.$$

This is less than 1 as long as

$$\frac{v_C}{\bar{v}} \geq \frac{r + \xi}{r + \lambda + \xi}.$$

Further, we require $\overline{p} \geq \underline{p}$. This implies

$$\frac{v_C}{\bar{v}} \geq \frac{[(r + \lambda)(\bar{u} - u_C) - (r + \xi)u_C]}{(r + \lambda)(\bar{u} - u_C) - \xi u_C}.$$

Note that $\underline{p}$ is greater than zero as long as $v_C/\bar{v} < 1$, which is true by assumption.

The boundaries given in Proposition 7 are as illustrated below in Figure **??**. Note that compared with figure 3 the only difference is that the set of parameters sustaining compromise is smaller due to less incentive to compromise because of delay effects.