Policy Reform*

Avidit Acharya[†] Juan Ortner[†]
Stanford University Boston University

February 17, 2017

Abstract

We construct a model of policy reform in which two players continually search for Pareto improving policies. The players have imperfect control over the proposals that are considered. Inefficient gridlock takes place due to the difficulty in finding moderate policies. The reform process is path dependent, with early agreements determining long-run outcomes. The process may also be cyclical, as players alternate between being more and less accommodating. Our model provides a noncooperative foundation for the "Raiffa path", by which bargainers gradually approach the Pareto frontier.

KEYWORDS: collective search, bargaining, path dependence, cycling, endogenous status quo, Raiffa path, gridlock, delay, inefficiency.

*We are grateful to David Baron, Steven Callander and Bart Lipman, and seminar audiences at Stanford and Boston University for their feedback.

[†]Acharya: avidit@stanford.edu, Ortner: jortner@bu.edu.

1 Introduction

Policy reform is a complicated process. The agents involved typically have imperfect control over the scope and direction of reform. One reason for this is that they must search for ideas upon which to build reform proposals, and it is hard to anticipate which ideas this search process yields and when. In fact, it is not uncommon to have third parties like think-tanks, lobbyists, NGOs, and interest groups contributing their diverse views to the development of a reform proposal. To learn the full implications of the legislation that they put forward, parties often have to wait until the final reform bill is made public and evaluated by the same groups that contributed to its formation as well as other experts, the media, and nonpartisan organizations.[1]

In this paper, we develop a tractable model of policy reform that accounts for the limited control that players have over the development of reform proposals, and we use it to study the dynamics of policy reform. We start by describing the structure of our model, and then discuss its multiple interpretations. We consider a two player complete information search model played over $T < \infty$ periods. The set of feasible policies is the simplex $X = \{\mathbf{x} \in \mathbb{R}^2_+ : x_1 + x_2 \leq 1\}$. At each period t, player $i = 1, 2$ obtains a flow payoff equal to the coordinate x_i^t of the policy $\mathbf{x}^t = (x_1^t, x_2^t)$ that is in effect. The policy in place at the start of the game is $(0, 0)$. In each period, a new policy is drawn randomly from the set of policies that are Pareto improvements to the policy last period, and the players sequentially decide whether to approve or disapprove the draw. The previous period policy is replaced if and only if both players approve the change; otherwise, it remains in place. Players share a common discount factor $\delta < 1$. This process of policy reform allows players to continually search for step-by-step improvements over existing policies. Since we are primarily interested in the limiting case with $T \to \infty$, players have the opportunity to get arbitrarily close to the Pareto frontier.[2]

This model with randomly generated policies can be interpreted as a bargaining model in which it is difficult for players to calibrate their offers. This is a natural assumption when the issue over which players are bargaining is complex, and they have to wait to discover good ideas on how to improve existing agreements. Examples of such

[1] The Congressional Budget Office (CBO) is an example of a nonpartisan organization that evaluates the budgetary consequences of bills, but there are numerous others.

[2] Studying the limit as $T \to \infty$ of our finite horizon model, rather than the infinite horizon case, is an equilibrium selection device: while the infinite horizon game has a plethora of subgame perfect equilibria (SPE), the game with deadline T has an essentially unique SPE.

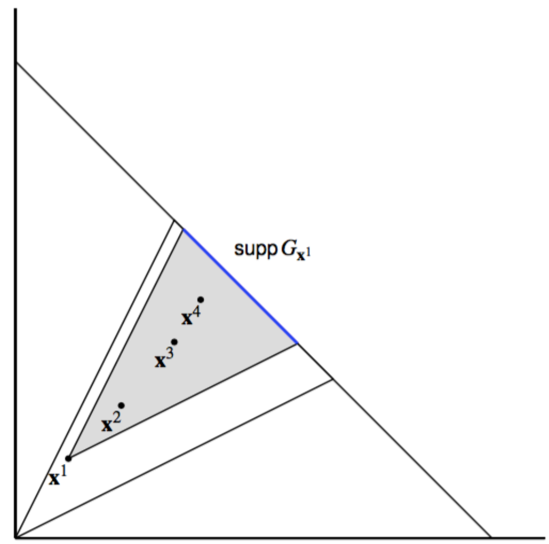

Figure 1: Path of play

complex issues include international climate change agreements, trade negotiations, and negotiations over international conflicts.

Our baseline model is one in which players have *no control* over the offer that is generated. In this sense, our model lies at the opposite extreme of the standard approach to bargaining theory (e.g. Rubinstein (1982) and Baron and Ferejohn (1989)) in which proposers have *full control* over the payoff-consequences of their proposals. We extend our baseline model to an intermediate case in which proposers have *partial control* over the payoff consequences of the offers they put on the table. We consider a setting in which, at each period, players probabilistically choose the distribution from which the policy will be drawn. We show that our main results carry through in this environment.

Our analysis delivers a clean equilibrium characterization. In any period, the set of policies that both players find acceptable is a cone defined by two lines with positive slope that pass through the last period's policy as its vertex. Figure 1 depicts such "acceptance cones" for a possible sequence of policies $\{\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3, \mathbf{x}^4, ...\}$ that are approved along the path of play. Policies that lie outside the acceptance cone are rejected even if they are Pareto superior to the status quo. The reason for this is that players cannot commit to approve future policies that disproportionally benefit their opponents. As a result, a player strictly prefers to reject Pareto superior offers that yield a substantial improvement to her opponent, but only a mild improvement for her, since she anticipates that approving such a policy will "close the door" in the future to many policies that she finds attractive. Since players discount the future, the periods of inaction produced by the rejection of Pareto improving policies generate inefficiency. This inaction and inefficiency are not produced by institutional constraints to reform, such as supermajority requirements, but simply by the difficulty in discovering moderate policies.[3]

As Figure 1 shows, a distinctive feature of our model is that players will typically reach a sequence of "interim" agreements, gradually approaching the Pareto frontier. Real world examples featuring such step-by-step bargaining dynamics include the nuclear disarmament negotiations between the US and the USSR, climate change agreements, and trade deals between the US and China.

The randomness of draws and the rigidity of the status quo in our model together imply that the policy reform process is path dependent. In any period the set of policies that players find acceptable depends on the current status quo. As a result, at each point in time the future path of play depends crucially on the policies that players agreed on at early stages of the reform process.

This path dependence disappears, however, when players become fully patient ($\delta \to 1$). In this case, the acceptance cone collapses to a line segment connecting the status quo to a point on the frontier: only policies on this line segment are implemented on the path of play. Intuitively, the cost in terms of forgone future payoff of implementing a policy that is more beneficial to one's opponent increases with δ. In the limit, the only policies that both players accept are those that give a payoff vector on this line segment. When policies are drawn from a symmetric distribution, the long run policy converges to an equal split of the surplus. In this case, the path that the equilibrium induces when players are arbitrarily patient coincides with the so-called "Raiffa path"; i.e., the path of policies proposed in Raiffa (1953) as a plausible outcome in settings in which the bargaining parties engage in step-by-step negotiations.[4]

[3]Our explanation for inefficient gridlock therefore stands in contrast to the explanations that have traditionally been highlighted by the literature, e.g. Brady and Volden (1998) and Krehbiel (2010).

[4]There have been other attempts to provide foundations for the Raiffa path. Livne (1989), Peters and Van Damme (1991), Diskin et al. (2011) and Samet (2009) provide axiomatizations for this bargaining solution. Myerson (2013) provides a noncooperative foundation by proposing a bargaining model with a finite deadline, in which players make offers with equal probability in each round. However, his model lacks the incremental improvements by which players approach the Pareto frontier since the unique equilibrium of his model has the players reach the frontier immediately. So his model provides a foundation for the Raiffa path only in the sense that as the number of periods goes to infinity, the agreement that the players reach in the first period converges to the long-run outcome (i.e., endpoint) of the Raiffa path.

Lastly, we show that our model may give rise to reform cycles, under which periods of high and low likelihood of agreements alternate. To understand why such cycles may occur, consider a game with a fixed deadline T. At the deadline, there is no benefit to waiting so players accept any policy that constitutes a Pareto improvement relative to the status quo. In period $T-1$, however, players are less accommodating, since they anticipate that the policy that they implement today will affect the policies that will be accepted tomorrow. Consider next period $T-2$. If the probability of drawing an acceptable policy at $T-1$ is small, the players know that they are unlikely to enact a reform in the next period, and, in all likelihood, will have to wait until the final period to change policy. Since waiting two periods is more costly than waiting one period, at time $T-2$ players are more accommodating than in period $T-1$. We provide conditions for such reform cycles to occur even in the limit as the deadline T goes to infinity.

Our paper is primarily related to the literature on collective search, e.g., Compte and Jehiel (2010), Albrecht et al. (2010) and Penn (2009). Compte and Jehiel (2010) and Albrecht et al. (2010) study models in which a group of agents sequentially sample alternatives from a distribution and have to choose when to stop. Closer to our model, Penn (2009) studies a setting with randomly generated alternatives and with an endogenous status-quo. Penn (2009) focuses on how the dynamic nature of the problem affects players' voting behavior among Pareto undominated policies. In contrast, our focus is on understanding the process by which policy approaches the Pareto frontier.[5]

Because the players in our model approach the Pareto frontier in incremental steps, our paper relates to prior work on incremental bargaining and partial agreements. Compte and Jehiel (2004) study a bargaining model in which each players' outside option depends on the history of offers. They show that, in this setting, players will make gradual concessions until they reach a final agreement. More recently, Acharya and Ortner (2013) analyze a model in which two players bargain over two issues, one of which will only be open for negotiation at a future date. The main result is that players may reach a partial agreement on the first issue, only to complete the agreement when the second issue becomes available.

Our result on commitment and inefficiency relates our paper to the literature on bargaining failures as a result of commitment problems; e.g., Fearon (1996), Powell (2004, 2006), Acemoglu and Robinson (2000, 2001), Ortner (2015). These papers focus

[5]The rigidity of the status quo relates our model to the growing literature on policy bargaining with an endogenous status quo – see, for instance, Kalandrakis (2004), Duggan and Kalandrakis (2012), Dziuda and Loeper (2015) and Bowen et al. (2013).

on understanding the conditions under which the players' inability to commit will result in bargaining inefficiencies. Instead, we focus on how the players' inability to commit shapes the evolution of policy towards the Pareto frontier.

Finally, our paper shares the spirit of Callander (2011), who also considers a setting of policy-making in complex environments. Callander (2011) focuses on how policy-makers learn about the payoff consequences of different policies from previous experiences. In contrast, bargainers in our model can fully evaluate the payoff consequences of the policies that they vote on. The complexity of the environment in our model is instead captured by the players' inability to draft policies that will deliver particular payoffs.

The paper proceeds as follows. Section 2 presents the model, establishes equilibrium existence and uniqueness, and provides a recursive characterization of equilibrium payoffs. Section 3 studies the main properties of our model in the limit as deadline T goes to infinity. Section 4 extends our baseline model to a setting in which players have partial control over offers. Section 5 concludes. All proofs are in the Appendix.

2 Model

2.1 The policy reform game

Two players, $i = 1, 2$, play the following *policy reform game*. Time is discrete, with an infinite horizon, and indexed by $t = 0, 1, 2, ...$ A *policy* is a pair

$$\mathbf{x} = (x_1, x_2) \in X := \{(y_1, y_2) \in \mathbb{R}^2_+ : y_1 + y_2 \leq 1\}.$$

In each of the first $T < \infty$ periods, players jointly decide whether to move policy from the current-period *status quo* $\mathbf{z}^t = (z_1^t, z_2^t) \in X$ to a new policy $\mathbf{x}$ drawn randomly from a distribution $F_{\mathbf{z}^t}$ with density $f_{\mathbf{z}^t}$ and full support over the set

$$X(\mathbf{z}^t) := \{\mathbf{x} \in X : x_i \geq z_i^t \text{ for } i = 1, 2\}$$

of Pareto superior policies to $\mathbf{z}^t$. We assume that, for all $\mathbf{z} \in X$ and all $\mathbf{x} \in X(\mathbf{z})$, $f_{\mathbf{z}}(\mathbf{x}) \in [\underline{f}, \overline{f}]$ for some constants $\overline{f} > \underline{f} > 0$. After the new policy $\mathbf{x}$ is drawn, the two players sequentially decide whether or not to accept it. If both players accept it, then the policy in place in period t becomes the new policy, so $\mathbf{x}^t = \mathbf{x}$. Otherwise, the status quo is implemented, so $\mathbf{x}^t = \mathbf{z}^t$. The next period's status quo is the previous period policy, so $\mathbf{z}^{t+1} = \mathbf{x}^t$ with $\mathbf{z}^0 = (0, 0)$. For all periods $t \geq T + 1$ the players cannot change policy, so $\mathbf{x}^t = \mathbf{x}^T$. We refer to the final reform period T as the *deadline*, and we will be interested in studying the limiting case of $T \to \infty$.

Both players are expected utility maximizers and share a common discount factor $\delta < 1$. If $\mathbf{x}^t = (x_1^t, x_2^t) \in X$ is the policy in place in period t, then player i earns a flow payoff x_i^t at time t. Player i's payoff from a sequence of policies $\{\mathbf{x}^t\}_{t=1}^{\infty}$ is thus

$$U_i\left(\{\mathbf{x}^t\}\right) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t x_i^t.$$

This describes the policy reform game with deadline T. We focus on the subgame perfect equilibria (SPE) of this game.

Proposition 1. *The policy reform game with deadline T has an SPE. All SPE of the policy reform game with deadline T are payoff equivalent.*

2.2 Recursive equilibrium characterization

For any $\mathbf{z} \in X$ and any $\mathbf{x} \in X(\mathbf{z})$, let

$$P_{\mathbf{z}}(\mathbf{x}) := \left(\frac{x_1 - z_1}{1 - z_1 - z_2}, \frac{x_2 - z_2}{1 - z_1 - z_2} \right) \tag{1}$$

$P_{\mathbf{z}}$ is a mapping that projects points in $X(\mathbf{z})$ onto X. We make the following assumption on the distributions $F_{\mathbf{z}^t}$ from which policies are drawn.

Assumption 1. *For every policy $\mathbf{z} \in X$, the density $f_{\mathbf{z}}$ satisfies*

$$f_{\mathbf{z}}(\mathbf{x}) = f(P_{\mathbf{z}}(\mathbf{x})) \qquad \forall \mathbf{x} \in X(\mathbf{z})$$

where $f := f_{(0,0)}$ is the density from which policies are drawn at the start of the game.

Assumption 1 states that, for any $\mathbf{z} \in X$, the distribution $F_{\mathbf{z}}$ over $X(\mathbf{z})$ from which policies are drawn when the status quo is $\mathbf{z}$ is "identical" to the distribution $F_{(0,0)}$ over X from which policies are drawn at the start of the game. We maintain this assumption throughout the rest of the paper. Its main implication is that a subgame starting at period $t \leq T$ with status quo policy $\mathbf{z} \in X$ is strategically equivalent to a game with deadline $T - t$ starting at policy $\mathbf{z}^0 = (0, 0)$.

To formalize this, for any deadline T, time $t \leq T$ and policy $\mathbf{z} \in X$, let $V_i(\mathbf{z}, t; T)$ be the continuation payoff that player i obtains under an SPE at a subgame starting at period t when the status quo is $\mathbf{z}^t = \mathbf{z}$. Let $W_i(T) = V_i((0,0), 0; T)$ be player i's SPE payoff at the start of a game. Then, we have:

Lemma 1. *For all $t \leq T$ and all possible values of the status quo $\mathbf{z}^t = \mathbf{z} = (z_1, z_2) \in X$, the players' equilibrium payoffs satisfy*

$$V_i(\mathbf{z}, t; T) = z_i + (1 - z_1 - z_2) W_i(T - t) \text{ for } i = 1, 2. \tag{2}$$

When the status quo at time t is $\mathbf{z}$, player i's equilibrium payoff is equal to the flow payoff z_i, that the player is guaranteed to get forever (by the persistence of the status quo), plus the payoff $(1 - z_1 - z_2) W_i(T - t)$ that the player obtains from bargaining for $T - t$ periods over the remaining surplus of size $1 - z_1 - z_2$.

We use Lemma 1 to provide a recursive characterization of the players' equilibrium payoffs. Note first that, at the last period T, players accept any policy in $X(\mathbf{z}^T)$, where $\mathbf{z}^T$ is the status quo policy. Consider next a period $t < T$ at which the status quo policy is $\mathbf{z} = (z_1, z_2) \in X$. Then, player i approves a policy $\mathbf{x} = (x_1, x_2) \in X(\mathbf{z})$ only if

$$(1 - \delta) x_i + \delta V_i(\mathbf{x}, t + 1; T) \geq (1 - \delta) z_i + \delta V_i(\mathbf{z}, t + 1; T). \tag{3}$$

Let $W_i = W_i(T - t - 1)$. Then using (2) in both sides of (3) and rearranging, player i accepts policy $\mathbf{x}$ when the status quo is $\mathbf{z}$ only if $x_i \geq \ell_{i,\mathbf{z}}(x_{-i}|W_i)$, where

$$\ell_{i,\mathbf{z}}(x_{-i}|W_i) := z_i + \frac{\delta W_i}{1 - \delta W_i}(x_{-i} - z_{-i})$$

$\ell_{i,\mathbf{z}}(x_{-i}|W_i)$ is the line in (x_i, x_{-i})-space with slope $\delta W_i/(1 - \delta W_i)$ that passes through the status quo $\mathbf{z}$. Define

$$A_{i,\mathbf{z}}(W_i) := \left\{\mathbf{x} \in X(\mathbf{z}) : x_i \geq \ell_{i,\mathbf{z}}(x_{-i}|W_i)\right\}.$$

Then, for any pair of payoffs $\mathbf{W} = (W_1, W_2)$ and for any $\mathbf{z} \in X$, the set

$$A_{\mathbf{z}}(\mathbf{W}) := A_{1,\mathbf{z}}(W_1) \cap A_{2,\mathbf{z}}(W_2) \tag{4}$$

is the set of policy draws that are accepted by both players at period $t < T$ when the status quo policy is $\mathbf{z}$ and $(W_1(T - t - 1), W_2(T - t - 1)) = (W_1, W_2)$. When $1 >$

$\delta(W_1+W_2)$, the line $\ell_{1,\mathbf{z}}(x_2|W_1)$ has larger slope greater than $\ell_{2,\mathbf{z}}(x_1|W_2)$ in (x_1,x_2)-space and $A_{\mathbf{z}}(\mathbf{W})$ is a cone with vertex $\mathbf{z}$. For any pair of values $\mathbf{W}$ we let $A(\mathbf{W}) := A_{(0,0)}(\mathbf{W})$ be the cone with vertex $(0,0)$. Such a cone is depicted in Figure 2.

For any integer $T>0$, let $\mathbf{W}(T) = (W_1(T), W_2(T))$ be the players' SPE payoffs in a game with deadline T. By our arguments above, a policy draw is accepted at the initial period if and only if it is in the set $A(\mathbf{W}(T-1))$ with $\mathbf{W}(T-1) = (W_1(T-1), W_2(T-1))$. Therefore, player i's payoff at the start of the game is

$$\begin{aligned} W_i(T) &= \mathrm{prob}(\mathbf{x} \in A(\mathbf{W}(T-1)))\mathbb{E}[(1-\delta)x_i + \delta V_i(\mathbf{x}, 1; T)|\mathbf{x} \in A(\mathbf{W}(T-1))] \\ &\qquad + \mathrm{prob}(\mathbf{x} \notin A(\mathbf{W}(T))[(1-\delta)0 + \delta V_i((0,0), 1; T)] \\ &= \mathrm{prob}(\mathbf{x} \in A(\mathbf{W}(T-1)))\mathbb{E}[x_i - (x_1+x_2)\delta W_i(T-1)|\mathbf{x} \in A(\mathbf{W}(T-1))] + \delta W_i(T-1), \end{aligned}$$

where the second line follows from equation (2).

Define the operator $\Phi = (\Phi_1, \Phi_2) : X \to X$, where for every payoff pair $\mathbf{W} = (W_1, W_2) \in X$ and for $i = 1, 2$,

$$\Phi_i(\mathbf{W}) := \mathrm{prob}(\mathbf{x} \in A(\mathbf{W}))\mathbb{E}[x_i - (x_1+x_2)\delta W_i|\mathbf{x} \in A(\mathbf{W})] + \delta W_i, \tag{5}$$

Let $\Phi^t(\mathbf{W})$ denote the t-th iteration of operator Φ over the pair $\mathbf{W} = (W_1, W_2)$.

Proposition 2. *In a policy reform game with deadline T,*

(i) the players' equilibrium payoffs satisfy $\mathbf{W}(T) = \Phi^{T+1}((0,0))$, and

(ii) the set of policies that are accepted by both players in any period $t \leq T$ is $A_{\mathbf{z}^t}(\mathbf{W}(T-t-1))$ where $\mathbf{z}^t$ is the status quo policy in period t and $\mathbf{W}(T-t-1)$ are the players' equilibrium payoffs in the policy reform game with deadline $T-t-1$.

Figure 2 plots the acceptance region $A(\mathbf{W})$ at the initial period of the game. As the figure shows, policies that constitute a Pareto improvement over the initial policy $(0,0)$ and that lie outside $A(\mathbf{W})$ are rejected, leading to inefficient outcomes.

The commitment problem lies at the heart of these inefficiencies. To see why, suppose that in period 0 policy $\mathbf{x} > (0,0)$ in Figure 2 is drawn. Policy $\mathbf{x}$ Pareto dominates the initial policy, but if $\mathbf{x}$ were to be implemented, then starting in period 1 the set of policies $A_{\mathbf{x}}(\mathbf{W})$ that both players accept would be the area inside the dashed lines in Figure 2. These policies are significantly worse for player 2 than the policies that could be implemented in the future if the status quo $(0,0)$ remains in place. So player 2

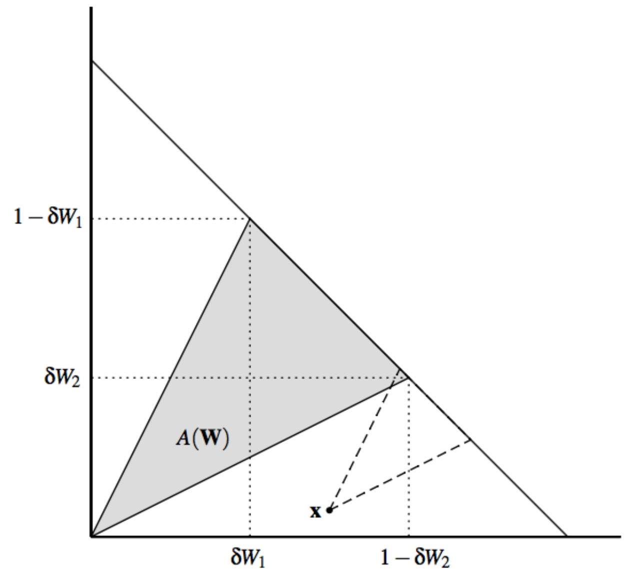

Figure 2: Acceptance region $A(\mathbf{W})$.

strictly prefers to maintain policy $(0,0)$ than to implement $\mathbf{x}$. Player 2 would approve policy $\mathbf{x}$ if player 1 could commit to accepting policies that are beneficial for player 2 in the future, bringing the trajectory of policy reform back towards the center. However, player 2 rightly anticipates that player 1 would not accept such policies in the future if policy $\mathbf{x}$ were to be implemented today.

3 Infinite horizon limit

Throughout this section, we study the properties of the equilibrium in the limit as T approaches ∞.

Definition 1. *We say that the equilibrium is* convergent *if the sequence* $\{\mathbf{W}(T)\} = \{\Phi^T(\mathbf{0})\}$ *converges as* $T \to \infty$. *Otherwise, we say that the equilibrium is* cycling.

Section 3.1 studies conditions under which equilibrium is convergent. Section 3.2 discusses some properties of convergent equilibria. Finally, Section 3.3 provides conditions under which equilibrium is cycling.

3.1 Conditions for convergence

The iterative characterization of equilibrium payoffs in Proposition 2 suggests that if the sequence of payoffs $\{\mathbf{W}(T)\}$ converges in T, then the limit is a fixed point of Φ. This is confirmed by the following lemma.

Lemma 2. *(i) Φ has a fixed point, and*

(ii) if the sequence of payoffs $\{\mathbf{W}(T)\}$ converges to $\mathbf{W}$, then $\mathbf{W}$ is a fixed point of Φ.

Our next result presents sufficient conditions for equilibrium to be convergent. In particular, it shows that equilibrium is convergent whenever players are patient enough.

Proposition 3. *There exists a threshold $\overline{\delta} < 1$ such that if $\delta > \overline{\delta}$ the equilibrium is convergent.*

Symmetric distributions. We now study conditions under which payoffs converge for the special case where F is *symmetric* about the 45° line, i.e. when its density f satisfies $f(x_1, x_2) = f(x_2, x_1)$ for all $(x_1, x_2) \in X$.

We start by noting that, when F is symmetric, both players get the same equilibrium payoffs: for all T, $\mathbf{W}(T) = (W_1(T), W_2(T))$ is such that $W_1(T) = W_2(T) =: W(T)$. A formal proof of this statement is given in Lemma A.2 in the Appendix.

For all T let $\hat{W}(T) = 2W(T)$ be the sum of the players' equilibrium payoffs in a game with deadline T. With a slight abuse of notation, let $A(\hat{W})$ be the acceptance region when $\mathbf{W} = (\hat{W}/2, \hat{W}/2)$. We define the operator $\Psi : [0,1] \to [0,1]$ as follows: for all $\hat{W}$,

$$\begin{aligned}\Psi(\hat{W}) &:= \Phi_1((\hat{W}/2, \hat{W}/2)) + \Phi_2((\hat{W}/2, \hat{W}/2)) \\ &= \text{prob}(\mathbf{x} \in A(\hat{W}))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A(\hat{W})](1 - \delta\hat{W}) + \delta\hat{W} \end{aligned} \tag{6}$$

It then follows from Proposition 2 that when F is symmetric, $\hat{W}(T) = \Psi^{T+1}(0)$.

Our next result provides a sufficient condition for equilibrium to be convergent in the especial case in which F is symmetric.

Proposition 4. *Suppose F is symmetric. Then, if $\Psi'(\hat{W}) > -1$ for all $\hat{W} \in [0,1]$, the equilibrium is convergent.*

To get a better sense as to when the condition in Proposition 4 holds, define $H(\hat{W}) :=$ $\text{prob}(\mathbf{x} \in A(\hat{W}))\mathbb{E}[x_1 + x_2|\mathbf{x} \in A(\hat{W})]$, so

$$\Psi'(\hat{W}) = \delta(1 - H(\hat{W})) + H'(\hat{W})(1 - \delta\hat{W}).$$

Note that $H'(\hat{W}) < 0$, and that the magnitude of this derivative depends on how much mass the distribution F puts on the boundary of the acceptance set: $|H'(\hat{W})|$ is large when F puts significant mass on the boundary of $A(\hat{W})$. Since $\delta(1 - H(\hat{W})) \geq 0$, the condition in Proposition 4 holds whenever the distribution F is sufficiently "dispersed."

Example 1. Assume F is a uniform distribution over X. In this case, for any $\hat{W} \in [0, 1]$,

$$\Psi(\hat{W}) = \delta\hat{W} + \frac{2}{3}(1 - \delta\hat{W})^2.$$

Note that $\Psi'(\hat{W}) = \frac{1}{3}(-1 + 4\delta\hat{W}) > -1$, so by Proposition 4 the equilibrium is convergent. Payoffs $\mathbf{W}(T) = (W_1(T), W_2(T))$ converge to $\mathbf{W} = (W_1, W_2)$, where for $i = 1, 2$,

$$W_i = \frac{1}{8\delta^2}(3 + \delta - \sqrt{9 + 6\delta - 15\delta^2}).$$

We note that, as $\delta \to 1$, equilibrium payoffs $\mathbf{W}$ converge to $(1/2, 1/2)$.

3.2 Properties of convergent equilibria

In this section, we assume that the equilibrium is convergent, so $\mathbf{W}(T)$ converges to some $\mathbf{W} = \Phi(\mathbf{W})$. We derive several properties of the limiting equilibrium.

We start by noting that, when $\mathbf{W}(T)$ converges to some $\mathbf{W} = (W_1, W_2) \in X$, each acceptance set $A_{\mathbf{x}^\tau}(\mathbf{W})$ is a cone with vertex $\mathbf{x}^\tau$, defined by two lines with slopes $\delta W_1/(1 - \delta W_1)$ and $(1 - \delta W_2)/\delta W_2$ that pass through the vertex (see Figure 2). This means that in the infinite horizon limit, the lines defining all of the acceptance cones are parallel, so the acceptance cones are nested.

Lemma 3. *(nested acceptance cones) Let $\{\mathbf{x}^t\}_{t=0}^{\infty}$ be a sequence of equilibrium policies. Then,*

$$A_{\mathbf{x}^0}(\mathbf{W}) \supseteq A_{\mathbf{x}^1}(\mathbf{W}) \supseteq A_{\mathbf{x}^2}(\mathbf{W}) \supseteq ...$$

Lemma 3 implies that there exist policies that are acceptable at some period t, but become no longer acceptable at period $t + 1$ despite also being Pareto improvements relative to the $t+1$ status quo policy $\mathbf{z}^{t+1}$.[6] Players don't implement some policy reforms that they would have previously accepted.

Long-run outcomes. The game essentially ends when players implement a policy $\mathbf{x} \in X$ on the Pareto frontier: if policy $\mathbf{x} \in X$ with $x_1 + x_2 = 1$ is implemented at time t, then $\mathbf{x}^\tau = \mathbf{x}$ for all periods $\tau \geq t$. We therefore call a policy $\mathbf{x}$ on the Pareto frontier a *long-run outcome* of the game. The game's unique equilibrium induces a distribution G over long-run outcomes; i.e., over points on the frontier. For any subgame starting with status-quo policy $\mathbf{z} \in X$, the continuation equilibrium at that subgame induces a distribution $G_\mathbf{z}$ over long-run outcomes. The next result summarizes some notable features of convergent equilibria, including that the distribution over long run outcomes changes along the path of play, and exhibits path dependence.

Proposition 5. *Suppose the equilibrium is convergent. Then,*

(i) (long run distribution) for any $\mathbf{z} \in X$, the support of distribution $G_\mathbf{z}$ is $\mathsf{supp}\, G_\mathbf{z} = \{\mathbf{y} \in X : y_1 + y_2 = 1\} \cap A_\mathbf{z}(\mathbf{W})$;

(ii) (path dependence) $G_\mathbf{z} \neq G_{\mathbf{z}'}$ for all $\mathbf{z}' \neq \mathbf{z}$;

(iii) (gradual certainty) For every sequence of equilibrium policies $\{\mathbf{x}^\tau\}_{\tau=0}^\infty$, $\mathsf{supp}\, G_{\mathbf{x}^{\tau+1}} \subseteq \mathsf{supp}\, G_{\mathbf{x}^\tau}$, with strict inclusion whenever $\mathbf{x}^{\tau+1} \neq \mathbf{x}^\tau$.

In the first period, any policy $\mathbf{x}$ on the Pareto frontier with $x_1 \in [\delta W_1, 1 - \delta W_2]$ lies in the support of $G = G_{(0,0)}$. As play progresses and the players implement policies that are closer to the Pareto frontier, the support of the long-run distribution shrinks. Figure 1 shows the support of $G_{\mathbf{x}^1}$ for some policy $\mathbf{x}^1$ on the path of play.

Patient players and the Raiffa path. We now study equilibrium behavior when players become arbitrarily patient; i.e., when $\delta \to 1$. We note that, by Proposition 3, the equilibrium is convergent whenever δ is larger than some threshold $\overline{\delta}$.

For each $\delta \in (\overline{\delta}, 1)$, we let $\mathbf{W}^\delta = (W_1^\delta, W_2^\delta)$ denote the players' limiting payoffs as $T \to \infty$ in a game with discount factor δ. We let G^δ denote the distribution over long run outcomes in the limiting equilibrium with discount factor δ.

Proposition 6. *Fix a sequence $\{\delta_n\} \to 1$, and a corresponding sequence of equilibrium payoffs $\{\mathbf{W}^{\delta_n}\}$. Then,*

[6] Formally, there exist policies $\mathbf{x} > \mathbf{x}^\tau$ such that $\mathbf{x} \in A_{\mathbf{x}^t}(\mathbf{W}) \backslash A_{\mathbf{x}^\tau}(\mathbf{W})$ for $\tau > t$.

(i) (generalized Raiffa path) $\lim_{n\to\infty} A(\mathbf{W}^{\delta_n}) = \{\mathbf{x} \in X : x_1/x_2 = W_1^*/W_2^*\}$;

(ii) (determinism) G^{δ_n} *converges to a dirac measure on* $(W_1^*, W_2^*) := \lim_{n\to\infty}(W_1^{\delta_n}, W_2^{\delta_n})$;

(iii) (efficiency) $\lim_{n\to\infty} W_1^{\delta_n} + W_2^{\delta_n} = 1$.

Moreover, if F *is symmetric,* $W_1^* = W_2^* = 1/2$.

Proposition 6(i) says that, as $\delta \to 1$, the set of policies that both players find acceptable converges to the line segment connecting $(0,0)$ and the point (W_1^*, W_2^*). Intuitively, the cost in terms of forgone future payoff of implementing a policy that is more beneficial to your opponent increases with δ. In the limit, the only policies that both players accept are those that give both players a payoff on this line segment. This implies that, as players become arbitrarily patient, there is no path dependence. Proposition 6(ii) says that as $\delta \to 1$ the path of play approaches deterministically a particular long run outcome, namely the players' equilibrium payoff split. Lastly, Proposition 6(iii) shows that the inefficiency of delay vanishes as players become infinitely patient. This occurs in spite of the fact that, as $\delta \to 1$, the acceptance region $A(\mathbf{W}^\delta)$ converges to a straight line, and so the probability of changing the policy in any given period goes to zero.

In general, the long-run agreement (W_1^*, W_2^*) depends on the distribution F. Proposition 6 establishes that in the special case in which the distribution is symmetric, both players obtain the same payoff, so $(W_1^*, W_2^*) = (1/2, 1/2)$.

As a result, when F is symmetric, the path of play that our model induces in the limit as $\delta \to 1$ is closely related to the sequential bargaining solution proposed by Raiffa (1953). Indeed, in our framework, Raiffa's sequential bargaining solution is the segment with slope 1 that connects the origin with the point $(1/2, 1/2)$ on the Pareto frontier.[7]

3.3 Cycling equilibrium

We now turn to cycling equilibria. We start by providing intuition as to why the equilibrium may be cycling. Players in our model trade off implementing a Pareto improving policy today against the benefit of waiting to see if they can move policy in a more preferred direction tomorrow. At the deadline T, there is no benefit to waiting so the players accept any policy in $X(\mathbf{z}^T)$. In the second to last period, however, players are less

[7]More generally, Raiffa's bargaining solution is the segment connecting the disagreement payoff with the Pareto frontier, and passing through the *utopia* payoff vector; i.e., the payoff vector that would result if each player obtained her preferred outcome. In our environment, the utopia payoff vector is $(1, 1)$.

accommodating, since they anticipate that the set of acceptable policies tomorrow will depend on the policy they implement today. Graphically, the acceptance cone becomes smaller (narrower) at period $T-1$.

Consider next period $T-2$. If the probability of changing the policy next period is sufficiently small (i.e., if distribution F places little mass on the acceptance cone tomorrow), players know that they are unlikely to enact a reform in the next period, and, in all likelihood, will have to wait until the final period to change policy. Since waiting for two periods is more costly than waiting only one period, players are more accommodating in period $T-2$ than in period $T-1$.

The arguments above suggests that payoffs $\mathbf{W}(T)$ may cycle for small values of T. We now provide conditions for the equilibrium to be cycling in the limit as $T \to \infty$.

For simplicity, we focus on the case in which the distribution F is symmetric. Recall from the discussion in Section 3.1 that when F is symmetric, the players have the same equilibrium payoffs and the sum of these payoffs is the $(T+1)$-th iteration over 0 of the operator Ψ defined in equation (6).

Proposition 7. *If F is symmetric then Ψ has a unique fixed point $\hat{W}^*$. If, in addition,*

(i) $\Psi(\hat{W}) \neq \hat{W}^*$ for all $\hat{W} \neq \hat{W}^*$, and

(ii) there exists $\varepsilon > 0$ such that $\Psi'(\hat{W}) \leq -1$ for all $\hat{W} \in [\hat{W}^* - \varepsilon, \hat{W}^* + \varepsilon]$,

then the equilibrium is cycling.

Under the conditions in Proposition 7, the players' equilibrium payoffs $\hat{W}(\tau)/2$ cycle around $\hat{W}^*/2$. Note that, in the symmetric case, the acceptance region $A_{\mathbf{z}}(\hat{W})$ is a cone with vertex $\mathbf{z}$ and lines with slopes $\frac{1-\delta\hat{W}/2}{\delta\hat{W}/2}$ and $\frac{\delta\hat{W}/2}{1-\delta\hat{W}/2}$. Therefore, the fact that payoffs $\hat{W}(\tau)/2$ cycle around $\hat{W}^*/2$ implies that there will be an alternation between periods of high likelihood of agreement and periods of low likelihood of agreement; i.e., the equilibrium features reform cycles.

We stress that these reform cycles in our model are not the result of structural changes in the environment, but rather a result of self-fulfilling changes in the players' expectations of how the game will be played in the future. In the context of policy reform, such cycles were observed by historians such as Schlesinger (1949), who emphasizes that the cycles we observe in the reform process frequently cannot be explained by changes in the fundamentals. Smith (1985) summarizes their view, writing that

"Even the advocates of the cycle of reform model have been hard pressed to identify the dynamic that drives the cycle. The most common explanation is organic, that society, like an animal, has a natural alteration between periods of rest and action..."

4 Strategic search

Our model, with random proposals, is intended to capture complexities in the environment that make it difficult for players to gauge the payoff consequences of their proposals. In this section, we present a natural extension of our framework in which players have some ability to influence the direction in which they will search for new policies.

From a bargaining perspective, our baseline model can be interpreted as a bargaining model in which the proposer has *no control* over the offer that is generated. In this sense, our model lies at the opposite extreme of the standard approach to bargaining theory (e.g. Rubinstein (1982) and Baron and Ferejohn (1989)) in which proposers have *full control* over the proposals that are considered. The extension we present in this section bridges the gap between the traditional approach and our baseline model by allowing proposers to have *partial control* over the payoff consequences of the offers they put on the table. We briefly describe the model here. A formal treatment of this extension appears in Appendix B.

Two players, $i = 1, 2$, play the following game. Time is discrete and indexed by $t = 0, 1, 2, ...$ The set of policies is X, and players have the same preferences over policies as in our baseline model. At each period $t = 0, 1, ..., T$, player $i = 1, 2$ is recognized with probability $1/2$. The recognized player chooses a distribution F from a finite set of distributions $\mathcal{F}_{\mathbf{z}^t}$, where $\mathbf{z}^t$ is the status quo policy. We assume that each distribution $F \in \mathcal{F}_{\mathbf{z}}$ has support in $X(\mathbf{z})$ and density f such that, for all $\mathbf{x} \in X(\mathbf{z})$, $f(\mathbf{x}) \in [\underline{f}, \overline{f}]$ for some $\overline{f} > \underline{f} > 0$.

After the new policy $\mathbf{x}$ is drawn, the two players sequentially decide whether or not to accept it. If both players accept it, then the policy in place in period t becomes the new policy, so $\mathbf{x}^t = \mathbf{x}$. Otherwise, the status quo is implemented, so $\mathbf{x}^t = \mathbf{z}^t$. The status quo at time $t + 1$ is the previous period policy, so $\mathbf{z}^{t+1} = \mathbf{x}^t$. For all periods $t \geq T + 1$ the players cannot change policy, so $\mathbf{x}^t = \mathbf{x}^T$. As in our baseline model, for any deadline T, this game has an essentially unique equilibrium which can be found by backward induction.

We make the following assumptions about the sets of distributions $\mathcal{F}_{\mathbf{x}}$. First, we assume that, for all $\mathbf{x}, \mathbf{y} \in X$, $\text{card}(\mathcal{F}_{\mathbf{x}}) = \text{card}(\mathcal{F}_{\mathbf{y}})$; i.e., all the sets $\mathcal{F}_{\mathbf{x}}$ have the same cardinality. Second, for all $\mathbf{x} \in X$ and all $F_{\mathbf{x}} \in \mathcal{F}_{\mathbf{x}}$ with density $f_{\mathbf{x}}$, there exists $F \in \mathcal{F} = \mathcal{F}_{(0,0)}$ with density f such that $f_{\mathbf{x}}(\mathbf{y}) = f(P_{\mathbf{x}}(\mathbf{y}))$ for all $\mathbf{y} \in X(\mathbf{x})$. Note that this assumption is a generalization of Assumption 1 to the current environment.

In Appendix B we show that this extended model retains the key features of our baseline model.

5 Conclusion

We have developed a model of policy reform built on the assumption that players have imperfect control over the proposals that are considered.

Our model suggests a new source of inefficient gridlock in the reform process, namely the difficulty in finding moderate policies that are acceptable to both of the players that are involved. In our model, the reform process is path dependent and may be cyclical as the players alternate between periods in which they are accommodating and stubborn. Inefficiency is driven by the commitment problem, and cycling is driven by an alternating pattern of changes in the players' self-fulfilling expectations about the likelihood of enacting a reform.

Our model provides an answer to the question of how two bargaining players approach the Pareto frontier. Our answer is that they do so in steps, while ensuring that these steps fit within the set of trajectories that ensure long-run moderation. We show that as the players become infinitely forward looking, the only acceptable trajectory is the one hypothesized by Raiffa under which the players are guaranteed to reach an equal division of the surplus.

Appendix

A Proofs

A.1 Proofs for Section 2

Proof of Proposition 1. In the subgame starting in period T, it is optimal for both players to accept any policy in $X(\mathbf{z}^T)$. Moreover, the policy that is drawn will remain in place in all future periods so the payoff to each player i at this subgame is

$$V_i(\mathbf{z}^T, T; T) = \mathbb{E}_{\mathbf{z}^T}[x_i] \tag{7}$$

where $\mathbb{E}_{\mathbf{z}^T}[\cdot]$ is the expectation operator under the distribution $F_{\mathbf{z}^T}$.

At any subgame starting in period $T-1$ with status quo policy z^{T-1}, it is optimal for player i to accept a policy $\mathbf{x} \in X(z^{T-1})$ if

$$(1-\delta)x_i + \delta V_i(\mathbf{x}, T; T) \geq (1-\delta)z_i^{T-1} + \delta V_i(\mathbf{z}^{T-1}, T; T) \tag{8}$$

So the set of policies that are acceptable to both players is

$$A_{\mathbf{z}^{T-1}} := \{\mathbf{x} \in X(\mathbf{z}^{T-1}) : (8) \text{ holds for both } i = 1, 2\}$$

This defines the payoff that each player i gets at such a subgame, which is

$$\begin{aligned} V_i(\mathbf{z}^{T-1}, T-1; T) = {} & \text{prob}(\mathbf{x} \in A_{\mathbf{z}^{T-1}})\mathbb{E}_{\mathbf{z}^{T-1}}\left[(1-\delta)x_i + \delta V_i(\mathbf{x}, T; T) \,|\mathbf{x} \in A_{\mathbf{z}^{T-1}}\right] \\ & + \text{prob}(\mathbf{x} \notin A_{\mathbf{z}^{T-1}})\left[(1-\delta)z_i^{T-1} + \delta V_i(\mathbf{x}^{T-1}, T; T)\right] \end{aligned}$$

Repeating these arguments for all $t < T$ establishes existence of a SPE, and uniqueness of SPE payoffs. ■

Proof of Lemma 1. Recall that for all $\mathbf{z} \in X$, $\mathbb{E}_{\mathbf{z}}[\cdot]$ is the expectation operator under distribution $F_{\mathbf{z}}$. Let $\mathbb{E}[\cdot]$ be the expectation operator under distribution $F_{(0,0)} = F$. We prove the result by induction.

Consider first a subgame starting at period $t = T$ with status quo $\mathbf{z}^T = \mathbf{z} \in X$. Note that

$$V_i(\mathbf{z}, T; T) = \mathbb{E}_{\mathbf{z}}[x_i] = z_i + (1 - z_1 - z_2)\mathbb{E}[x_i],$$

where the first equality follows from equation (7) and the second equality follows from Assumption 1.

Now, consider the policy reform game with deadline $T = 0$. By equation (7), player i's equilibrium payoffs satisfy $W_i(0) = \mathbb{E}[x_i]$. Hence,

$$V_i(\mathbf{z}, T; T) = z_i + (1 - z_i - z_j)W_i(0)$$

which establishes the basis case.

For the induction step, suppose that (2) holds for all t such that $T - t = 0, 1, ..., n - 1$ and for all $\mathbf{z} \in X$. Fix a subgame starting at period $\tilde{t}$ with $T - \tilde{t} = n$ and with status quo $\mathbf{z}^{\tilde{t}} = \mathbf{z} \in X$. Let $A_{\mathbf{z}}(\tilde{t})$ be the set of policies that both players accept at period $\tilde{t}$ when $\mathbf{z}^{\tilde{t}} = \mathbf{z}$; that is,

$$\begin{aligned} A_{\mathbf{z}}(\tilde{t}) &= \left\{\mathbf{x} \in X(\mathbf{z}) : (1-\delta)x_i + \delta V_i(\mathbf{x}, \tilde{t}+1; T) \geq (1-\delta)z_i + \delta V_i(\mathbf{z}, \tilde{t}+1; T) \text{ for } i = 1,2\right\} \\ &= \left\{\mathbf{x} \in X(\mathbf{z}) : (x_i - z_i) \geq (x_1 + x_2 - z_1 + z_2)\delta W_i(T - \tilde{t} - 1) \text{ for } i = 1,2\right\}, \end{aligned}$$

where the second line follows since, by the induction hypothesis, (2) holds for $t = \tilde{t} + 1$. Note then that

$$\begin{aligned} V_i(\mathbf{z}, \tilde{t}; T) = {} & \text{prob}(\mathbf{x} \in A_{\mathbf{z}}(\tilde{t}))\mathbb{E}_{\mathbf{z}}\left[(1-\delta)x_i + \delta V_i(\mathbf{x}, \tilde{t}+1; T) \,\middle|\, \mathbf{x} \in A_{\mathbf{z}}(\tilde{t})\right] \\ & + \text{prob}(\mathbf{x} \notin A_{\mathbf{z}}(\tilde{t}))\left((1-\delta)z_i + \delta V_i(\mathbf{z}, \tilde{t}+1; T)\right) \\ = {} & \text{prob}(\mathbf{x} \in A_{\mathbf{z}}(\tilde{t}))\mathbb{E}_{\mathbf{z}}\left[x_i + (1 - x_1 - x_2)\delta W_i(T - \tilde{t} - 1) \,\middle|\, \mathbf{z} \in A_{\mathbf{z}}(\tilde{t})\right] \\ & + \text{prob}(\mathbf{x} \notin A_{\mathbf{z}}(\tilde{t}))\left(z_i + (1 - z_1 - z_2)\delta W_i(T - \tilde{t} - 1)\right) \\ = {} & \text{prob}(\mathbf{x} \in A_{\mathbf{z}}(\tilde{t}))\mathbb{E}_{\mathbf{z}}\left[(x_i - z_i) + (z_1 + z_2 - x_1 - x_2)\delta W_i(T - \tilde{t} - 1) \,\middle|\, \mathbf{x} \in A_{\mathbf{z}}(\tilde{t})\right] \\ & + z_i + (1 - z_1 - z_2)\delta W_i(T - \tilde{t} - 1) \qquad (9) \end{aligned}$$

where the second equality follows since, by the induction hypothesis, (2) holds for $t = \tilde{t} + 1$, and the last inequality follows since $\text{prob}(\mathbf{x} \notin A_{\mathbf{z}}(\tilde{t})) = 1 - \text{prob}(\mathbf{x} \in A_{\mathbf{z}}(\tilde{t}))$.

Consider next a game with deadline $T - \tilde{t}$. Let $\tilde{A}$ be the set of policies that both players accept at the first period of the game:

$$\begin{aligned}\tilde{A} &= \left\{\mathbf{x} \in X : (1-\delta)x_i + \delta V_i(\mathbf{x}, 1; T - \tilde{t}) \geq \delta V_i((0,0), 1; T - \tilde{t}) \text{ for } i = 1, 2\right\} \\ &= \left\{\mathbf{x} \in X : x_i \geq (x_1 + x_2)\delta W_i(T - \tilde{t} - 1) \text{ for } i = 1, 2\right\},\end{aligned}$$

where the second line follows since, by the induction hypothesis, for all $V_i(\mathbf{x}, 1; T - \tilde{t}) = x_i + (1 - x_i - x_j)W_i(T - \tilde{t})$ for all $\mathbf{x}$. Player i's payoff in this game is equal to

$$\begin{aligned}W_i(T - \tilde{t}) &= \text{prob}(\mathbf{x} \in \tilde{A})\mathbb{E}\left[(1-\delta)x_i + \delta V_i(\mathbf{x}, 1; T - \tilde{t}) \,\middle|\, \mathbf{x} \in \tilde{A}\right] + \text{prob}(\mathbf{x} \notin \tilde{A})\delta V_i((0,0), 1; T - \tilde{t}) \\ &= \text{prob}(\mathbf{x} \in \tilde{A})\mathbb{E}\left[x_i - (x_1 + x_2)\delta W_i(T - \tilde{t} - 1) \,\middle|\, \mathbf{x} \in \tilde{A}\right] + \delta W_i(T - \tilde{t} - 1)\end{aligned} \tag{10}$$

Assumption 1 implies that

$$\begin{aligned}&\text{prob}(\mathbf{x} \in A_{\mathbf{z}}(\tilde{t}))\mathbb{E}_{\mathbf{z}}\left[x_i - z_i + (z_1 + z_2 - x_1 - x_2)\delta W_i(T - \tilde{t} - 1) \,\middle|\, \mathbf{x} \in A_{\mathbf{z}}(\tilde{t})\right] \\ =&(1 - z_1 - z_2)\text{prob}(\mathbf{x} \in \tilde{A})\mathbb{E}\left[x_i - (x_1 + x_2)\delta W_i(T - \tilde{t} - 1) \,\middle|\, \mathbf{x} \in \tilde{A}\right].\end{aligned}$$

Combining this with (9) and (10),

$$V_i(\mathbf{z}, \tilde{t}; T) = z_i + (1 - z_1 - z_2)W_i(T - \tilde{t}).$$

which establishes the result. ■

Proof of Proposition 2. (i) The proof is by induction. Consider the game with deadline $T = 0$. Since it is optimal for both players to accept any policy $\mathbf{x} \in X$ that is drawn, player i's payoff in this game satisfies $W_i(T) = \mathbb{E}[x_i] = \Phi_i((0,0))$.

Suppose next that $W_i(\tau) = \Phi_i^{\tau+1}((0,0))$ for all $\tau = 0, ..., T-1$, and consider game with deadline T. The set of policies that both players accept in the initial period are given by

$$\begin{aligned}\tilde{A} &= \{\mathbf{x} \in X : (1-\delta)x_i + \delta V_i(\mathbf{x}, 1; T) \geq \delta V_i((0,0), 1; T) \text{ for } i = 1, 2\} \\ &= \{\mathbf{x} \in X : x_i \geq (x_1 + x_2)\delta W_i(T - 1) \text{ for } i = 1, 2\},\end{aligned}$$

where the second line follows from Lemma 1. Player i's payoff $W_i(T)$ satisfies

$$
\begin{aligned}
W_i(T) &= \text{prob}(\mathbf{x} \in \tilde{A})\mathbb{E}\left[(1-\delta)x_i + \delta V_i(\mathbf{x}, 1; T) \,\middle|\, \mathbf{x} \in \tilde{A}\right] + \text{prob}(\mathbf{x} \notin \tilde{A})\delta V_i((0,0), 1; T) \\
&= \text{prob}(\mathbf{x} \in \tilde{A})\mathbb{E}\left[x_i - (x_1 + x_2)\delta W_i(T-1) \,\middle|\, \mathbf{x} \in \tilde{A}\right] + \delta W_i(T-1) \qquad (11)
\end{aligned}
$$

where the equality follows after using Lemma 1. By the induction hypothesis, $\mathbf{W}(T-1) = \Phi^T((0,0))$, and so $\tilde{A} = A(\Phi^T((0,0)))$. Using this in (11), it follows that $W_i(T) = \Phi(\Phi^T((0,0))) = \Phi^{T+1}((0,0))$.

(ii) Fix a period $t \leq T$ and a policy $\mathbf{z} \in X$, and consider a subgame starting at period t with status quo policy $\mathbf{z}^t = \mathbf{z}$. At such a subgame, player i finds it optimal to accept policies $\mathbf{x} \in X(\mathbf{z})$ satisfying

$$(1-\delta)x_i + \delta V_i(\mathbf{x}, t+1; T) \geq (1-\delta)z_i + \delta V_i(\mathbf{z}, t+1; T)$$

or, using equation (2) in Lemma 1, policies that satisfy

$$x_i - z_i \geq (x_1 + x_2 - z_1 - z_2)\delta W_i(T-t-1). \qquad (12)$$

The set of policies that both players accept at period t when the status quo is $\mathbf{z}^t = \mathbf{z}$ is therefore the set of policies $\mathbf{x} \in X(\mathbf{z})$ for which (12) is satisfied for both $i = 1, 2$. This is precisely the set $A_{\mathbf{z}}(\mathbf{W}(T-t-1))$ of policies defined in (4). ■

Proof of Lemma 2. (i) Φ is continuous and maps X onto itself, so by Brouwer's fixed point theorem, it has a fixed point.

(ii) If $\{\mathbf{W}(T)\}$ converges to $\mathbf{W}$, then Proposition 2(i) implies that

$$\mathbf{W} = \lim_{T\to\infty} \Phi^T((0,0)) = \Phi\big(\lim_{T\to\infty} \Phi^{T-1}((0,0))\big) = \Phi(\mathbf{W}),$$

so $\mathbf{W}$ is a fixed-point of Φ. ■

For every $\delta < 1$, let $A^\delta(\mathbf{W})$ and Φ^δ be, respectively, the acceptance sets and the operator defined in equation (5) when the discount factor is δ. Let $\mathbf{W}^\delta = (W_1^\delta, W_2^\delta)$ be a fixed point of Φ^δ: for $i, j = 1, 2, i \neq j$,

$$\begin{aligned} W_i^\delta &= \delta W_i^\delta + \text{prob}(\mathbf{x} \in A^\delta(\mathbf{W}^\delta))\mathbb{E}[x_i - (x_i + x_j)\delta W_i^\delta | \mathbf{x} \in A^\delta(\mathbf{W}^\delta)] \\ \iff W_i^\delta &= \frac{\text{prob}(\mathbf{x} \in A^\delta(\mathbf{W}^\delta))\mathbb{E}[x_i | \mathbf{x} \in A(\mathbf{W}^\delta)]}{1 - \delta + \delta \text{prob}(\mathbf{x} \in A^\delta(\mathbf{W}^\delta))\mathbb{E}[x_i + x_j | \mathbf{x} \in A^\delta(\mathbf{W}^\delta)]}. \end{aligned}$$

Then,

$$W_1^\delta + W_2^\delta = \frac{\text{prob}(\mathbf{x} \in A^\delta(\mathbf{W}^\delta))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A^\delta(\mathbf{W}^\delta)]}{1 - \delta + \delta \text{prob}(\mathbf{x} \in A^\delta(\mathbf{W}^\delta))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A^\delta(\mathbf{W}^\delta)]}. \tag{13}$$

Lemma A.1. *Fix a sequence of discount factors $\{\delta_n\} \to 1$, and let $\mathbf{W}^{\delta_n} = (W_1^{\delta_n}, W_2^{\delta_n}) \in X$ be a sequence such that $\mathbf{W}^{\delta_n} = \Phi^{\delta_n}(\mathbf{W}^{\delta_n})$ for all n. Then, $\lim_{n\to\infty}(W_1^{\delta_n} + W_2^{\delta_n}) = 1$.*

Proof. Towards a contradiction, suppose this is not true. Hence, there exists a sequence $\{\delta_n\} \to 1$ and a positive number $\eta > 0$ such that $W_1^{\delta_n} + W_2^{\delta_n} < 1 - \eta$ for all n. Note that this implies that $A^{\delta_n}(\mathbf{W}^{\delta_n})$ has a non-empty interior for all n. Since the distribution F has density f such that $f(\mathbf{x}) \geq \underline{f} > 0$ for all $\mathbf{x}$, there exists a constant $K > 0$ such that $\text{prob}(\mathbf{x} \in A^{\delta_n}(\mathbf{W}^{\delta_n})) > K$ for all n. It follows that

$$\lim_{n\to\infty} W_1^{\delta_n} + W_2^{\delta_n} = \lim_{n\to\infty} \frac{\text{prob}(\mathbf{x} \in A^{\delta_n}(\mathbf{W}^{\delta_n}))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A^{\delta_n}(\mathbf{W}^{\delta_n})]}{1 - \delta_n + \delta_n \text{prob}(\mathbf{x} \in A^{\delta_n}(\mathbf{W}^{\delta_n}))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A^{\delta_n}(\mathbf{W}^{\delta_n})]} = 1,$$

a contradiction. Hence, it must be that $W_1^{\delta_n} + W_2^{\delta_n} \to 1$ as $\delta_n \to 1$. ■

Proof of Proposition 3. We start by showing that, for any $\delta < 1$, there exists $\underline{V}^\delta < 1$ with $\delta \underline{V}^\delta \to 1$ as $\delta \to 1$ such that, for all $\mathbf{W} = (W_1, W_2)$ with $W_1 + W_2 < \underline{V}^\delta$, $\Phi_1(\mathbf{W}) + \Phi_2(\mathbf{W}) > W_1 + W_2$. Note that this property implies that, for any fixed point $\mathbf{W}^\delta = (W_1^\delta, W_2^\delta)$ of Φ^δ, it must be that $\underline{V}^\delta \leq W_1^\delta + W_2^\delta$.[8]

To see why such a $\underline{V}^\delta$ exists, pick $\underline{g} \in (0, \underline{f})$ with $\underline{g} < 1$ and note that for any $\mathbf{W} \in X$,

$$\begin{aligned} \Phi_1^\delta(\mathbf{W}) + \Phi_2^\delta(\mathbf{W}) &= \delta(W_1 + W_2) + \text{prob}(x \in A^\delta(\mathbf{W}))\mathbb{E}[x_1 + x_2 | x \in A^\delta(\mathbf{W})](1 - \delta(W_1 + W_2)) \\ &\geq \delta(W_1 + W_2) + \frac{1}{3}\underline{f}(1 - \delta(W_1 + W_2))^2 \\ &> \delta(W_1 + W_2) + \frac{1}{3}\underline{g}(1 - \delta(W_1 + W_2))^2, \end{aligned} \tag{14}$$

[8]If $\underline{V}^\delta > W_1^\delta + W_2^\delta$, then $\Phi_1^\delta(\mathbf{W}^\delta) + \Phi_2^\delta(\mathbf{W}^\delta) > W_1^\delta + W_2^\delta$, contradicting the fact that $\mathbf{W}^\delta$ is a fixed point of Φ^δ.

where the first inequality follows since $f(\mathbf{x}) \geq \underline{f} > 0$ for all $\mathbf{x}$.[9] Equation (14) implies that $\Phi_1^\delta(\mathbf{W}) + \Phi_2^\delta(\mathbf{W}) > W_1 + W_2$ for all $\mathbf{W}$ when

$$\frac{1}{3}\underline{g}\frac{(1-\delta(W_1+W_2))^2}{1-\delta} > W_1 + W_2.$$

Let $\underline{V}^\delta$ be the smallest solution to $\frac{1}{3}\underline{g}\frac{(1-\delta\underline{V}^\delta)^2}{1-\delta} = \underline{V}^\delta$; i.e.,

$$\underline{V}^\delta = \frac{3(1-\delta)}{2\underline{g}\delta^2}\left(1 + \frac{2\underline{g}\delta}{3(1-\delta)} - \sqrt{1 + \frac{4\underline{g}\delta}{3(1-\delta)}}\right).$$

It follows that $\Phi_1(\mathbf{W}) + \Phi_2(\mathbf{W}) > W_1 + W_2$ for all $\mathbf{W}$ with $W_1 + W_2 < \underline{V}^\delta$. Note that $\underline{V}^\delta < 1$ for $\delta < 1$, and that $\delta\underline{V}^\delta \to 1$ as $\delta \to 1$.

We show next that there exists $\underline{\delta} < 1$ such that, for all $\delta > \underline{\delta}$ and for all $\mathbf{W} = (W_1, W_2) \in X$ with $W_1 + W_2 \geq \underline{V}^\delta$, $(\Phi^\delta)^T(\mathbf{W})$ converges to a fixed point of Φ^δ as $T \to \infty$. Towards establishing this, note that for $i, j = 1, 2$, $i \neq j$,

$$\begin{aligned}
\frac{\partial\Phi_i^\delta(\mathbf{W})}{\partial W_i} &= \delta - \delta\int_{\mathbf{x}\in A^\delta(\mathbf{W})}(x_1+x_2)f(\mathbf{x})d\mathbf{x} \in \left[\delta - \frac{\overline{f}}{3}\delta(1-\delta(W_1+W_2)), \delta\right]\\
\frac{\partial\Phi_i^\delta(\mathbf{W})}{\partial W_j} &= -\int_0^{1-\delta W_j}\delta x_i^2 f\left(x_i, \frac{\delta W_j x_i}{1-\delta W_j}\right)dx_i\frac{1-\delta(W_1+W_2)}{(1-\delta W_j)^3} \in \left[-\frac{\overline{f}}{3}\delta(1-\delta(W_1+W_2)), 0\right],
\end{aligned}$$

where we used the assumption that $f(\mathbf{x}) \leq \overline{f}$ for all $\mathbf{x} \in X$. Since $\delta\underline{V}^\delta \to 1$ as $\delta \to 1$, there exists $\underline{\delta} < 1$ such that, for all $\delta > \underline{\delta}$ and all $\mathbf{W} \in X$ with $W_1 + W_2 \geq \delta\underline{V}^\delta$, and $\delta - \frac{2\overline{f}}{3}\delta(1-\delta(W_1+W_2)) \geq 0$. Note that, for all $\delta > \underline{\delta}$ and all $\mathbf{W} \in X$ with $W_1+W_2 \geq \delta\underline{V}^\delta$

$$\frac{\partial\Phi_i^\delta(\mathbf{W})}{\partial W_i} + \frac{\partial\Phi_i^\delta(\mathbf{W})}{\partial W_j} \in (-\delta, \delta). \tag{15}$$

Fix $\delta > \underline{\delta}$, and let $Y^\delta := \{\mathbf{W} \in X : W_1 + W_2 \geq \delta\underline{V}^\delta\}$. Let $||\cdot||$ be the sup-norm on $\mathbb{R}^2$. By equation (15), for all $\mathbf{W}, \mathbf{W}' \in Y^\delta$ and for $i = 1, 2$, $|\Phi_i(\mathbf{W}) - \Phi_i(\mathbf{W}')| \leq \delta \times ||\mathbf{W} - \mathbf{W}'||$. Hence, for all $\mathbf{W}, \mathbf{W}' \in Y^\delta$, $||\Phi(\mathbf{W}) - \Phi(\mathbf{W}')|| \leq \delta \times ||\mathbf{W} - \mathbf{W}'||$.

[9] For all $\mathbf{W} \in X$,

$$\begin{aligned}
\text{prob}(x \in A^\delta(\mathbf{W}))\mathbb{E}[x_1 + x_2 | x \in A^\delta(\mathbf{W})] &= \int_{\mathbf{x}\in A^\delta(\mathbf{W})}(x_1+x_2)f(\mathbf{x})d\mathbf{x}\\
&\geq \underline{f}\int_{\mathbf{x}\in A^\delta(\mathbf{W})}(x_1+x_2)d\mathbf{x} = \frac{1}{3}\underline{f}(1-\delta(W_1+W_2)).
\end{aligned}$$

Moreover, $\Phi(\mathbf{W}) \in Y^\delta$ for all $\mathbf{W} \in Y^\delta$.[10] Note that this implies that, for all $\delta > \underline{\delta}$ and for all $\mathbf{W} \in Y^\delta$, $(\Phi^\delta)^T(\mathbf{W})$ converges to a fixed point of Φ^δ as $T \to \infty$.

Lastly, we show that the equilibrium is convergent whenever $\delta > \underline{\delta}$. Fix $\delta > \underline{\delta}$. There are two cases to consider: (i) $\Phi^\delta(\mathbf{0}) \in Y^\delta$, and (ii) $\Phi^\delta(\mathbf{0}) \notin Y^\delta$. In case (i), for any deadline $T \geq 0$, $\mathbf{W}(T) = (\Phi^\delta)^T(\Phi^\delta(\mathbf{0}))$ converges to a fixed point of Φ^δ as $T \to \infty$.

Consider next case (ii). Since $\Phi_1^\delta(\mathbf{W}) + \Phi_2(\mathbf{W}) > W_1 + W_2$ for all $\mathbf{W} = (W_1, W_2)$ with $W_1 + W_2 < \underline{V}^\delta$, there exists $t \geq 1$ such that $\Phi_1^\delta((\Phi^\delta)^t(\mathbf{0})) + \Phi_2((\Phi^\delta)^t(\mathbf{0})) \geq \underline{V}^\delta$. Hence, by our arguments above, $(\Phi^\delta)^{t+s}(\mathbf{0})$ converges to a fixed point of Φ^δ as $s \to \infty$, and so the equilibrium is convergent. ■

Lemma A.2. *If F is symmetric then the players have the same equilibrium payoffs for all deadlines, i.e. $W_1(T) = W_2(T) =: W(T)$ for all $T \geq 0$.*

Proof. If F is symmetric, then

$$W_1(0) = \Phi_1((0,0)) = \mathbb{E}[x_1] = \mathbb{E}[x_2] = \Phi_2((0,0)) = W_2(0)$$

Now suppose that $W_1(t) = W_2(t)$ for all $t = 0, ..., T-1$. Then, $W_1(T-1) = W_2(T-1)$ implies that the set $A((W_1(T-1), W_2(T-1))$ is symmetric, i.e. if $\mathbf{x} = (x_1, x_2) \in A((W_1(T-1), W_2(T-1))$ then $(x_2, x_1) \in A((W_1(T-1), W_2(T-1))$. Then, we have

$$\begin{aligned}
W_1(T) &= \Phi_1(\mathbf{W}(T-1)) \\
&= \text{prob}(\mathbf{x} \in A(\mathbf{W}(T-1))\mathbb{E}[x_1 - (x_1+x_2)W_1(T-1)|\mathbf{x} \in A(\mathbf{W}(T-1))] + \delta W_1(T-1) \\
&= \text{prob}(\mathbf{x} \in A(\mathbf{W}(T-1))\mathbb{E}[x_2 - (x_1+x_2)W_2(T-1)|\mathbf{x} \in A(\mathbf{W}(T-1))] + \delta W_2(T-1) \\
&= \Phi_2(\mathbf{W}(T-1)) = W_2(T),
\end{aligned}$$

[10] Proof: Note that the function $G(V) = \delta V + \frac{1}{3}\underline{g}(1-\delta V)^2$ is increasing in V whenever $\delta - \frac{2}{3}\overline{g}\delta(1-\delta V) \geq 0$. Then, since $\delta - \frac{2}{3}\delta\underline{g}(1-\delta(W_1+W_2)) > \delta - \frac{2}{3}\overline{f}\delta(1-\delta(W_1+W_2)) \geq 0$ for all $\delta \geq \underline{\delta}$ and all $\mathbf{W} \in Y^\delta$, it follows that,

$$\begin{aligned}
\Phi_1^\delta(\mathbf{W}) + \Phi_2^\delta(\mathbf{W}) &\geq \delta(W_1+W_2) + \frac{1}{3}\underline{g}(1-\delta(W_1+W_2))^2 \\
&\geq \delta\underline{V}^\delta + \frac{1}{3}\underline{g}(1-\delta\underline{V}^\delta)^2 = \underline{V}^\delta,
\end{aligned}$$

for all $\delta \geq \underline{\delta}$ and all $\mathbf{W} \in Y^\delta$. Hence, for all $\delta \geq \underline{\delta}$ and all $\mathbf{W} \in Y^\delta$, $\Phi(\mathbf{W}) \in Y^\delta$.

where the third equality follows since $W_1(T-1) = W_2(T-1)$ and since F is symmetric. ■

Proof of Proposition 4. For any $\hat{W} \in [0,1]$, define

$$H(\hat{W}) := \text{prob}(\mathbf{x} \in A(\hat{W}))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A(\hat{W})],$$

so that $\Psi(\hat{W}) = \delta\hat{W} + H(\hat{W})(1-\delta\hat{W})$. Note that $H'(\hat{W}) \leq 0$: indeed, $\hat{W}'' > \hat{W}'$ implies that $A(\hat{W}'') \subset A(\hat{W}')$, so for any $\hat{W}'' > \hat{W}'$,

$$\text{prob}(\mathbf{x} \in A(\hat{W}''))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A(\hat{W}'')] \leq \text{prob}(\mathbf{x} \in A(\hat{W}'))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A(\hat{W}')].$$

It then follows that $\Psi'(\hat{W}) = \delta(1 - H(\hat{W})) + H'(\hat{W})(1 - \delta\hat{W}) \leq \delta < 1$ for all $\hat{W} \in [0,1]$. When $\Psi'(\hat{W}) > -1$ for all $\hat{W} \in [0,1]$, $|\Psi'(\hat{W})| < 1$ for all $\hat{W} \in [0,1]$. This implies that Ψ is a contraction, and the sequence $\{\hat{W}(T)\}$ converges to its unique fixed point. Hence, the equilibrium is convergent. ■

A.2 Proofs for Section 3.2

Proof of Lemma 3. Fix any $\tau \geq t$. Since $\mathbf{x}^{\tau+1} \in A_{\mathbf{x}^\tau}(\mathbf{W})$ we have

$$x_i^{\tau+1} \geq \ell_{i,\mathbf{x}^\tau}(x_{-i}^{\tau+1}|W_i) = x_i^\tau + \frac{\delta W_i}{1-\delta W_i}(x_{-i}^{\tau+1} - x_{-i}^\tau)$$

for both $i = 1,2$. For any $\mathbf{y} = (y_1, y_2) \in A_{\mathbf{x}^{\tau+1}}(\mathbf{W})$, add $y_{-i}\delta W_i/(1-\delta W_i)$ to both sides of the above inequality and rearrange to get

$$x_i^{\tau+1} + \frac{\delta W_i}{1-\delta W_i}(y_{-i} - x_{-i}^{\tau+1}) \geq x_i^\tau + \frac{\delta W_i}{1-\delta W_i}(y_{-i} - x_{-i}^\tau)$$

This means that

$$\ell_{i,\mathbf{x}^{\tau+1}}(y_{-i}|W_i) \geq \ell_{i,\mathbf{x}^\tau}(y_{-i}|W_i), \qquad i = 1,2. \tag{16}$$

Thus if $\mathbf{y} \in A_{\mathbf{x}^{\tau+1}}(\mathbf{W})$ then $y_i \geq \ell_{i,\mathbf{x}^{\tau+1}}(y_{-i}|W_{-i})$ for $i = 1,2$, and by (16), $y_i \geq \ell_{i,\mathbf{x}^\tau}(y_{-i}|W_i)$ for $i = 1,2$. This means that $\mathbf{y} \in A_{\mathbf{x}^\tau}(\mathbf{W})$, and thus $A_{\mathbf{x}^{\tau+1}}(\mathbf{W}) \subseteq A_{\mathbf{x}^\tau}(\mathbf{W})$. ■

Proof of Proposition 5. For each $\mathbf{z} \in X$, define $LR_{\mathbf{z}} := A_{\mathbf{z}}(\mathbf{W}) \cap \{\mathbf{y} \in X : y_1 + y_2 = 1\}$. Since distribution $F_{\mathbf{z}}$ has full support and since $LR_{\mathbf{z}} \subseteq A_{\mathbf{z}}(\mathbf{W})$, any point in $LR_{\mathbf{z}}$ can arise as a long term outcome; i.e., $LR_{\mathbf{z}} \subseteq \mathsf{supp}\, G_{\mathbf{z}}$.

Consider next a subgame starting at period t with $\mathbf{z}^t = \mathbf{z}$. By Lemma 3, $\mathbf{x}^\tau \in A_{\mathbf{z}}(\mathbf{W})$ for all $\tau \geq t$. Since $LR_{\mathbf{z}} = A_{\mathbf{z}}(\mathbf{W}) \cap \{\mathbf{z} \in X : y_1 + y_2 = 1\}$, any point on the frontier that is not in $LR_{\mathbf{z}}$ cannot arise as a long term outcome when $\mathbf{z}^t = \mathbf{z}$. Hence, $\mathsf{supp}\, G_{\mathbf{z}} \subseteq LR_{\mathbf{z}}$.

This establishes that $\mathsf{supp}\, G_{\mathbf{z}} = LR_{\mathbf{z}}$, and it follows that $G_{\mathbf{z}} \neq G_{\mathbf{z}'}$ for $\mathbf{z} \neq \mathbf{z}'$. Lemma 3 then implies that along a realized equilibrium path $\{x_\tau\}_{\tau=t}^\infty$, we have $\mathsf{supp}\, G_{\mathbf{x}^{\tau+1}} \subseteq \mathsf{supp}\, G_{\mathbf{x}^\tau}$. The inclusion is strict when $\mathbf{x}^{\tau+1} \neq \mathbf{x}^\tau$ since $LR_{\mathbf{x}^{\tau+1}} \neq LR_{\mathbf{x}^\tau}$ in this case. ■

Proof of Proposition 6. Fix a sequence $\{\delta_n\}$ with $\delta_n \to 1$. For each n, let $\mathbf{W}^{\delta_n} = (W_1^{\delta_n}, W_2^{\delta_n})$ be the players' equilibrium payoffs in the limit as $T \to \infty$ in a game with discount factor δ_n. By Lemma 2, for each n, $\mathbf{W}^{\delta_n}$ is a fixed point of Φ^{δ_n}. By Lemma A.1, $\{\mathbf{W}^{\delta_n}\}$ is such that $\lim_{n\to\infty} W_1^{\delta_n} + W_2^{\delta_n} = 1$. This establishes part (iii).

Consider next part (ii). By Proposition 5, for each n the support of the long-run distribution G^{δ_n} is

$$A(\mathbf{W}) \cap \{\mathbf{y} \in X : y_1 + y_2 = 1\} = \{\mathbf{x} \in X : x_1 + x_2 = 1 \text{ and } x_1 \in [\delta W_1^{\delta_n}, 1 - \delta W_2^{\delta_n}]\}.$$

By part (iii), $\delta_n(W_1^{\delta_n} + W_2^{\delta_n})$ converges to 1 as $n \to \infty$. Hence, $[\delta_n W_1^{\delta_n}, 1 - \delta_n W_2^{\delta_n}]$ converges to a point W_1^*, and so G^{δ_n} converges to a dirac measure on (W_1^*, W_2^*).

Finally, recall that

$$A^{\delta_n}(\mathbf{W}^{\delta_n}) = \left\{\mathbf{x} \in X : x_i \geq \frac{\delta_n W_i^{\delta_n}}{1 - \delta_n W_i^{\delta_n}} x_{-i} \text{ for } i = 1, 2\right\}.$$

Using part (iii), $A^{\delta_n}(\mathbf{W}^{\delta_n})$ converges to $\{\mathbf{x} \in X : x_1/x_2 = W_1^*/W_2^*\}$. ■

Proof of Proposition 7. First we prove that if F is symmetric then the fixed point of Ψ is unique. Operator Ψ is continuous and maps $[0,1]$ onto itself, so by Brouwer's fixed point theorem, it has a fixed point.

Let $\hat{W}$ be a fixed point of Ψ. Then, $\hat{W}$ satisfies

$$\hat{W} = \frac{\text{prob}(\mathbf{x} \in A(\hat{W}))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A(\hat{W})]}{1 - \delta + \delta \text{prob}(\mathbf{x} \in A(\hat{W}))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A(\hat{W})]}. \tag{17}$$

Note that $A(\hat{W}'') \subset A(\hat{W}')$ for any $\hat{W}'' > \hat{W}'$. Therefore, for any $\hat{W}'' > \hat{W}'$,

$$\text{prob}(\mathbf{x} \in A(\hat{W}''))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A(\hat{W}'')] \leq \text{prob}(\mathbf{x} \in A(\hat{W}'))\mathbb{E}[x_1 + x_2 | \mathbf{x} \in A(\hat{W}')].$$

Thus, the right side of (17) is decreasing in $\hat{W}$, and so Ψ has a unique fixed point.

Next, the sum of the players' equilibrium payoff in a game with deadline T is $\hat{W}(T) = \Psi^{T+1}(0)$. By standard results in dynamical systems (e.g., Theorem 4.2 in De la Fuente (2000)), under conditions (i) and (ii) in the statement of the proposition the sequence $\{\hat{W}(T)\}$ does not converge. So the equilibrium must be cycling. ■

B Strategic search – not for publication

In this appendix we study the extension described in Section 4. We start by noting that this game has an essentially unique SPE – this can be established using the same arguments as in the proof of Proposition 1.

Fix a deadline T and a SPE σ^*. For every time $t \leq T$ and any $\mathbf{z} \in X$, let $V_i(\mathbf{z}, t; T)$ by player i's SPE continuation payoff at period t in a game with deadline T when the status quo policy at time t is $\mathbf{z}$. Let $W_i(T)$ denote player i's equilibrium payoff at the start of the game.

Recall that we made the following assumptions on the sets of distributions $\mathcal{F}_{\mathbf{x}}$. First, for all $\mathbf{x}, \mathbf{y} \in X$, $\text{card}(\mathcal{F}_{\mathbf{x}}) = \text{card}(\mathcal{F}_{\mathbf{y}})$; i.e., all the sets $\mathcal{F}_{\mathbf{x}}$ have the same cardinality. Second, for all $\mathbf{x} \in X$ and all $F_{\mathbf{x}} \in \mathcal{F}_{\mathbf{x}}$ with density $f_{\mathbf{x}}$, there exists $F \in \mathcal{F} = \mathcal{F}_{(0,0)}$ with density f such that $f_{\mathbf{x}}(\mathbf{y}) = f(P_{\mathbf{x}}(\mathbf{y}))$ for all $\mathbf{y} \in X(\mathbf{x})$. Note that these assumptions are a generalization of Assumption 1 to the current environment.

The following result generalizes Lemma 1 to the current environment. The proof is identical to the proof of Lemma 1, and hence omitted (the proof relies on the assumptions described above on the sets of distributions $\mathcal{F}_{\mathbf{x}}$).

Lemma B.1. *For all $t \leq T$ and all $\mathbf{z}^t = \mathbf{z} = (z_1, z_2) \in X$,*

$$V_i(\mathbf{z}, t; T) = z_i + (1 - z_1 - z_2)W_i(T - t). \tag{18}$$

Lemma B.1 can be used to obtain a recursive characterization of equilibrium payoffs. Consider a period $t \leq T$ at which the status quo policy is $\mathbf{z} = (z_1, z_2) \in X$. As in our baseline model, player i approves a policy $\mathbf{x} = (x_1, x_2) \in X(\mathbf{z})$ only if

$$(1-\delta)x_i + \delta V_i(\mathbf{x}, t+1; T) \geq (1-\delta)z_i + \delta V_i(\mathbf{z}, t+1; T)$$
$$x_i + (1 - x_1 - x_2)\delta W_i(T-t-1) \geq z_i - (1 - x_1 - x_2)\delta W_i(T-t-1),$$

where we used Lemma B.1. Let $W_i = W_i(T-t-1)$. Then, at period t player i accepts policy $\mathbf{x}$ when the status quo is $\mathbf{z}$ only if $x_i \in A_{i,\mathbf{z}}(W_i) = \{\mathbf{x} \in X(\mathbf{z}) : x_i \geq \ell_{i,\mathbf{z}}(x_{-i}|W_i)\}$, where $\ell_{i,\mathbf{z}}(x_{-i}|W_i)$ is defined as in the main text. For any pair of payoffs $\mathbf{W} = (W_1, W_2)$ and for any $\mathbf{z} \in X$, the set $A_{\mathbf{z}}(\mathbf{W})$ defined in the main text is the set of policy draws that are accepted by both players at period $t < T$ when the status quo policy is $\mathbf{z}$ and $(W_1(T-t-1), W_2(T-t-1)) = (W_1, W_2)$.

Consider a game with deadline T. Suppose player $i = 1, 2$ is recognized to choose the distribution from which the policy will be drawn at the initial period. If player i chooses distribution $F \in \mathcal{F}$, she obtains payoffs equal to

$$\text{prob}_F(x \in A(\mathbf{W}(T-1)))\mathbb{E}_F[x_i - (x_1 + x_2)\delta W_i(T-1)|\mathbf{x} \in A(\mathbf{W}(T-1))] + \delta W_i(T-1).$$

For any $\mathbf{W} \in X$ and for $i = 1, 2$, let

$$F^*_{\mathbf{W},i} \in \arg\max_{F \in \mathcal{F}} \text{prob}_F(x \in A(\mathbf{W}))\mathbb{E}_F[x_i - (x_1 + x_2)W_i|\mathbf{x} \in A(\mathbf{W})],$$

and let $F^*_{\mathbf{W}} := \frac{1}{2}F^*_{\mathbf{W},1} + \frac{1}{2}F^*_{\mathbf{W},2}$. Note that, when $\mathbf{W}(T-1) = \mathbf{W}$, the initial period policy is drawn from distribution $F^*_{\mathbf{W}}$.

Define the operator $\Gamma : X \to X$ as follows: for $i = 1, 2$ and for all $\mathbf{W} \in X$,

$$\Gamma_i(\mathbf{W}) = \text{prob}_{F^*_{\mathbf{W}}}(x \in A(\mathbf{W}))\mathbb{E}_{F^*_{\mathbf{W}}}[x_i - (x_1 + x_2)\delta W_i|\mathbf{x} \in A(\mathbf{W})] + \delta W_i.$$

For any integer t, let Γ^t denote the t-th iteration of operator Γ.

For any integer T, let $\mathbf{W}(T)$ denote the players' SPE payoffs in a game with deadline T. The following result extends Proposition 2 to the current environment – the proof uses the same arguments as the proof of Proposition 2, and hence we omit it.

Proposition B.1. *In the equilibrium of the policy reform game with endogenous proposals and deadline T,*

(i) the players' equilibrium values satisfy $\mathbf{W}(T) = \Gamma^{T+1}((0,0))$, and

(ii) the set of policies that are accepted by both players in any period $t \leq T$ is $A_{\mathbf{z}^t}(\mathbf{W}(T-t-1))$ where $\mathbf{z}^t$ is the status quo policy in period t and $\mathbf{W}(T-t-1)$ are the players' equilibrium payoffs in the policy reform game with deadline $T-t-1$.

This characterization of equilibrium payoffs can be used to generalize the main results in the main text to the current environment. First, the equilibrium features inefficient delays. Second, when the equilibrium is convergent, the acceptance regions are nested, and the distribution over long-run outcomes that the equilibrium induces at a subgame starting with status quo payoff $\mathbf{z}$ has support equal to $\{\mathbf{y} \in X : y_1 + y_2 = 1\} \cap A_{\mathbf{z}}(\mathbf{W})$. Therefore, the equilibrium also displays path-dependence. Lastly, it can be shown that Proposition 6 continues to hold in this setting, so the equilibrium outcome also becomes deterministic in the limit as $\delta \to 1$.

References

Acemoglu, D. and J. A. Robinson (2000): "Why Did the West Extend the Franchise? Democracy, Inequality, and Growth in Historical Perspective," *The Quarterly Journal of Economics*, 115, 1167–1199.

——— (2001): "A theory of political transitions," *American Economic Review*, 938–963.

Acharya, A. and J. Ortner (2013): "Delays and partial agreements in multi-issue bargaining," *Journal of Economic Theory*, 148, 2150–2163.

Albrecht, J., A. Anderson, and S. Vroman (2010): "Search by committee," *Journal of Economic Theory*, 145, 1386–1407.

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

Bowen, R., Y. Chen, and H. Eraslan (2013): "Mandatory Versus Discretionary Spending: the Status Quo Effect," *American Economic Review*, 104, 2941–2974.

Brady, D. W. and C. Volden (1998): *Revolving gridlock: Politics and policy from Carter to Clinton*, Westview Pr.

Callander, S. (2011): "Searching for good policies," *American Political Science Review*, 105, 643–662.

Compte, O. and P. Jehiel (2004): "Gradualism in Bargaining and Contribution Games," *Review of Economic Studies*, 71, 975–1000.

——— (2010): "Bargaining and majority rules: A collective search perspective," *Journal of Political Economy*, 118, 189–221.

De la Fuente, A. (2000): *Mathematical methods and models for economists*, Cambridge University Press.

Diskin, A., M. Koppel, and D. Samet (2011): "Generalized Raiffa solutions," *Games and Economic Behavior*, 73, 452–458.

Duggan, J. and T. Kalandrakis (2012): "Dynamic Legislative Policy Making," *Journal of Economic Theory*, 147, 1653–1688.

Dziuda, W. and A. Loeper (2015): "Dynamic Collective Choice with Endogenous Status Quo," *Journal of Political Economy*, forthcoming.

Fearon, J. D. (1996): "Bargaining over objects that influence future bargaining power," *Unpublished manuscript.*

Kalandrakis, T. (2004): "A Three-Player Dynamic Majoritarian Bargaining Game," *Journal of Economic Theory*, 116, 294–322.

Krehbiel, K. (2010): *Pivotal politics: A theory of US lawmaking*, University of Chicago Press.

Livne, Z. A. (1989): "On the status quo sets induced by the Raiffa solution to the two-person bargaining problem," *Mathematics of Operations Research*, 14, 688–692.

Myerson, R. B. (2013): *Game theory*, Harvard university press.

Ortner, J. (2015): "A theory of political gridlock," *Theoretical Economics.*

Penn, E. M. (2009): "A model of farsighted voting," *American Journal of Political Science*, 53, 36–54.

Peters, H. and E. Van Damme (1991): "Characterizing the Nash and Raiffa bargaining solutions by disagreement point axioms," *Mathematics of Operations Research*, 16, 447–461.

Powell, R. (2004): "The inefficient use of power: Costly conflict with complete information," *American Political Science Review*, 98, 231–241.

——— (2006): "War as a commitment problem," *International organization*, 60, 169.

Raiffa, H. (1953): "Arbitration schemes for generalized two-person games," *Annals of Mathematics Studies*, 28, 361–387.

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

Samet, D. (2009): "What if Achilles and the tortoise were to bargain? An argument against interim agreements," .

Schlesinger, A. M. (1949): *Paths to the Present*, Macmillan Co.

Smith, T. W. (1985): "Cycles of Reform? A Summary of Trends Since World War II," *Paper presented at the American Association for Public Opinion Research*.