The Politics of Kludges

Keiichi Kawai and Hongyi Li

Current Version: November 30, 2014
(Preliminary and Incomplete)

Abstract

This paper develops a model of policymaking under political conflict. In the model, policy changes may take the form of kludges: incremental modifications to existing policy that leave fundamental inefficiencies unresolved, resulting in excessively complex policies. Kludged policies emerge and persist when there is political conflict between ideologically opposed parties. Parties are more likely to implement policy kludges in the presence of frictions that impede policymaking. Further, political conflict may lead to obstructionist behavior, whereby one party deliberately introduces policy kludges to impede opponents' attempts to change policy.

1 Introduction

Complexity in public policy often arises from the use of kludges – piecemeal attempts to modify existing policy that paper over existing problems rather than resolving them in a fundamental way. Consider the Affordable Care Act (ACA), an American healthcare law passed in 2010. A primary goal of the ACA was to expand access to health insurance, and it did so by introducing mechanisms (including mandates, subsidies and insurance exchanges) that were designed to complement, rather than replace, the existing complex patchwork of private and public insurance options. A common view from both proponents and opponents was that the goals of the ACA could have been achieved, with far less policy complexity, by alternative policy solutions such as a single-payer healthcare system. Further, some parties who were sympathetic to the goals of the ACA nonetheless opposed its passage, because they believed that said passage would cement undesirable features of the existing insurance system and block any future move to a less-complex single-payer system.

This example highlights two key features of policy kludges. First, kludges allow policymakers to avoid difficult policy overhauls, but at the cost of additional policy complexity. Second, policy kludges exacerbate the persistence of existing policy.

In this paper, we present a dynamic model of policymaking that captures these features. In the model, policymaking is incremental: policy is composed of a sequence of rules that are added or removed, one rule at a time. The novel aspect of our model is that policy is backwards-dependent: when undoing policy, the policymaker has to remove recently-added rules before he can remove older rules. Our motivation for modeling backward-dependence is the idea that new rules build upon, and fill gaps in, existing rules. This complementarity implies backward-dependence: policy modifications rely on existing policy for coherence, and thus render removal of existing policy even more costly and difficult. One example is the Alternative Minimum Tax (AMT) of 1969: many observers consider the AMT to be an unnecessarily complex component of the American tax code, but also believe that it will be difficult to remove or drastically change the AMT because many (more recently enacted) aspects of the federal tax system have come to rely on the AMT. As Teles (2013) points out, "new ideas have to be layered over old programs rather than replace them ... "

A policymaker may seek to add or remove rules to achieve his policy ideals. Kludge has a natural interpretation in this setting, in terms of excessive complexity: a policy is kludged if an alternative policy achieves the same position with less complexity, i.e., using less rules. The fundamental tension that policymakers face in the model is a tradeoff between achieving policy ideals and reducing policy kludge. To fix ideas, consider a policymaker whose ideal policy position lies on the left of a left-right ideological spectrum. The policymaker may progress towards his ideal by adding new left-leaning rules to modify existing policy. Such additions patch over existing rules instead of removing them, which means that policy complexity may be exacerbated by a combination of old and new rules. Further, this complexity may persist in the long-run: backward-dependence means that post-modification, a policymaker cannot remove undesirable old rules without first undoing his own modifications, which he may be reluctant to do. Alternatively, the policymaker may choose to start off removing undesirable existing rules before implementing his preferred rules; in doing so, he reduces policy complexity, but may delay the attainment of his policy ideal.

Our analysis focuses on policymaking under political conflict. We consider a game between two policymakers, with conflicting ideological preferences, who take turns to make policy. This captures the premise that policymakers have to anticipate how opponents will respond in the future to their current policy changes. We provide a taxonomy of the factors that tend to favor the emergence and persistence of kludge.

First, kludge arises from conflict between parties with strong and conflicting ideological preferences. An ideologically zealous policymaker prioritizes adding new rules to patch over existing policy (and thus achieving his policy ideal quickly) over removing rules made by his opponent (and thus reducing complexity in the long-run). Over the long run, conflicting rules set by conflicting policymakers cancel out each others' effect on policy positions, while introducing additional compexity with each additional rule; the result is persistent policy kludge.

One implication of this logic is that an increase in political competition may lead to an increase in policy kludges. The model thus suggests that countries with intense political competition (such as the United States) may experience a greater degree of policy complexity than countries with little political competition (such as China or Singapore).

The model also suggests that kludge is more likely to arise in settings where institutional frictions make it difficult for policymakers to effect policy changes. This result provides some insight into recent claims about the origins of kludge in American public policy. An influential article by Teles (2013), while discussing political kludge in the context of American public policy, promulgates the common view that the excessive complexity of existing American public policy is driven by inherent conservatism of American governing institutions, which make it difficult to create new laws and undo existing laws.

A second phenomenon that emerges under political conflict is *obstructionism.* An ideologically zealous policymaker may intentionally introduce rules that do not improve the policy position, but serve to obstruct his opponent's future policy changes (which would obviously be unfavourable to the original policymaker). Interestingly, the optimal form of such an obstructionary strategy depends on the strength of the opponent's ideological preferences.

Against a ideologically moderate opponent, the policymaker generates *intentional complexity*: he introduces ideology-neutral rules that, due to backward-dependence, have to be removed before other ideology-relevant rules may subsequently be removed. One way to think about this result is that policymakers, in an attempt to protect their policy gains, may implement policy in a excessively complex fashion that stymies the undoing of said policy. This is consistent with the observation that policymakers often construct complicated bureaucracies (with no apparent ideological purpose) to implement policy (see, e.g., Moe 1989); our interpretation is that these bureaucracies serve as a moat to protect their policy gains from opponents.

On the other hand, against an ideologically zealous opponent, the policymaker may engage in *strategic extremism*: he pursues policy outcomes that are even more ideologically extreme than his preferences would naively dictate, so as to delay the progress of his opponent in the future. The model thus provides a potential explanation for strategic extremism in policymaking, and suggests that such behavior emerges in conflicts against zealous opponents.

Third, we show that the potential for political conflict may lead to what we call *appeasement.* In a situation where a moderate policymaker has the opportunity to add new rules that favour his preferred policy position, he may instead choose to do nothing. The reason is that the moderate party anticipates that his ideologically zealous opponent will (upon taking control) behave aggressively and patch over existing policy with new rules, rather than undoing any rules that the moderate may have implemented. This means that any rules that the moderate implements will not have any effect on policy in the long run, while irrevocably introducing unnecessary kludge. The moderate thus prefers to avoid such kludge by not introducing any new rules in the first place. In this sense, moderate policymakers may choose not to act even when they do not currently face any political constraints.

While we emphasize the applications of our model to public policy, we believe that it has relevance for rulemaking within organizations as well. In particular, given that the rulemaking process within organizations is incremental and backwards-dependent (as anyone who has ever sat on a university committee will surely attest to), the model suggests that intra-organizational conflict may result in the development of kludge, in the form of excessively complicated and bureaucratic rules.

1.1 Literature Review

Ely (2011) study how kludge may arise and persist in single-player adaptive processes. Their focus is on how random shocks to the environment may cause kludges to accumulate, resulting in persistently inefficient outcomes. We take a distinct approach to study kludge. First, we focus on how conflict between multiple players with conflicting ideals may exacerbate the formation of kludge. Second, instead of assuming that players behave adaptively (and thus myopically), we allow for patient and strategic behavior, but assume that players face exogenous political constraints and can only make small, 'local' changes to policy. This approach allows us to derive distinctive implications for the emergence of policy complexity under political conflict.

A number of papers from various literatures explore the idea that rule development may be path-dependent. Callander and Hummel (2014) consider a model where successive policymakers with conflicting preferences strategically experiment to find their preferred policy. The first policymaker benefits from a 'surprising' experiment outcome, because it deters experimentation by the second policymaker and thus preserves any policy gains by the first policymaker. Ellison and Holden (2013) study a model of endogenous rule development where there are exogenous constraints on the extent to which new rules may 'overwrite' old rules. Compared to these models, our paper introduces path dependence through a distinct mechanism – backwards dependence – and thus produces very different implications.

2 Model

A **policy** ϕ is an ordered set of rules. For any two policies $\phi = \{d_1, ...d_n\}$ and $\phi' = \{d'_1, ..., d'_{n'}\}$, denote the concatenated policy composed of ϕ followed by ϕ' as $\phi \sqcup \phi' = \{d_1, ..., d_n, d'_1, ..., d'_{n'}\}$. We say that ϕ is *adjacent* to ϕ' (equivalently, ϕ' is adjacent to ϕ) if $\phi' = \phi \sqcup \{d\}$ for some singleton rule d.

Each rule is characterized by its ideological **direction**: $d \in \{-1, 0, 1\}$. We may think of $d = -1$ as a left-leaning rule, $d = 1$ as a right-leaning rule, and $d = 0$ as a neutral rule. The ideological **position** $\rho(\phi)$ of policy ϕ is the sum of all rules in ϕ, and the **complexity** $\gamma(\phi)$ of policy ϕ is the total number of rules in ϕ. For example, if $\phi = \{1, -1, 0, 1\}$, then $\rho(\phi) = -1$ and $\gamma(\phi) = 4$ (see Figure 1).

Two players, (L)eft and (R)ight, play a policymaking game in continuous time. Denote the time-t policy as ϕ_t; the game starts with the empty policy, which we call the *origin*: $\phi_0 = \{\}$. In any time t, one of the two players $I_t \in \{L, R\}$ is in control. At each instant t, from the current policy ϕ_t, I_t can target any policy adjacent to ϕ_t.

1. While I_t targets a policy ϕ' that is an extension of ϕ_t (i.e. $\phi' = \phi_t \sqcup \{d\}$ for some singleton d), the policy randomly switches from ϕ_t to ϕ' with constant arrival rate p,

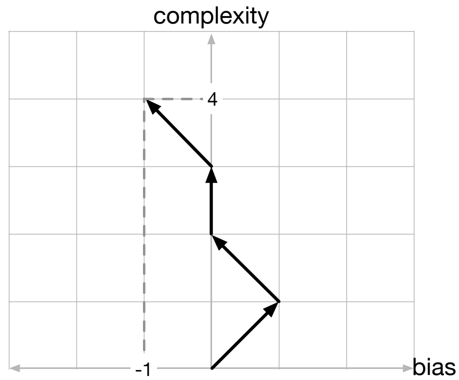

Figure 1: $\phi = (1, -1, 0, 1)$

and we say that I_t *extends* policy.

2. While I_t targets the (unique) policy ϕ'' that is a truncation of ϕ_t (i.e. $\phi_t = \phi'' \sqcup \{d\}$ for some singleton d), the policy randomly switches from ϕ_t to ϕ'' with constant arrival rate q, and we say that I_t *undoes* policy.
3. I_t can also choose not to target any adjacent policy, in which case the policy ϕ_t remains unchanged, and we say that I_t *stagnates*.

Notice that extending policy corresponds to adding a new rule to the existing policy, whereas undoing policy corresponds to removing the most recently-added rule from the existing policy. We may think of p and q as reflecting the magnitude of institutional frictions in the policymaking process; the larger p (resp. q) is, the more difficult for policymakers to add (resp. remove) rules. We assume that rules are easier to add than to remove: $p > q > 0$.

Change of control from one play to the other is stochastic. Player L starts the game in control. At each instant that L is in control, he loses control to R with constant arrival rate $\lambda > 0$. Once R gains control, he is in control forever after. We say that λ is player L's **vulnerability**.

Preferences The instantaneous payoff of player $I \in \{L, R\}$ at time t depends on the policy in place:

$$\pi_I(\phi_t) = -\zeta_I |\hat{\rho}_I - \rho(\phi_t)| - \gamma(\phi_t) \tag{1}$$

where $\rho_I^* \in \mathbb{Z}$ is his **ideal**, and ζ_I is his **ideological zeal**. With this payoff function, players prefer policies that are closer in ideological position to their ideals. Each player I discounts future payoff at rate r, i.e., his continuation payoff at t is

$$V_{I,t} = E\left[\int_{\tau=t}^{\infty} e^{-r\tau} \pi_I(\phi_\tau)\, d\tau\right]$$

The two players have conflicting ideological positions: $-\rho_L^* < 0$ and $\rho_R^* > 0$. Given the current policy ϕ_t, define player I's *favoured direction* $d_{I,t}$ to be the direction from the current policy ϕ_t's ideological position to I's ideal: $d_{I,t} = \frac{\hat{\rho}_I - \rho(\phi_t)}{|\hat{\rho}_I - \rho(\phi_t)|}$. (If $\hat{\rho}_I - \rho(\phi_t) = 0$, i.e. ϕ_t is an ideal policy, then $d_{I,t}$ is undefined.)

We restrict attention to $\zeta_L > 1$ and $\zeta_R > 1$, i.e. each player has a sufficiently strong preference over ideological position. This assumption ensures that adding rules is potentially profitable, so that there is a meaningful tradeoff between extending and undoing policy: by adding a rule in his favoured direction, player I increases his instantaneous payoff by $\zeta_I - 1 > 0$ (policy moves one step closer to I's ideal, while policy complexity increases by one). We'll think of players with ζ_I close to one as moderates, and players with high ζ_I as zealots.

For each $I \in \{L, R\}$, A policy ϕ is I-pure if there exists no other policy ϕ' that is both weakly closer to I's ideal $|\hat{\rho}_I - \rho(\phi')| \leq |\hat{\rho}_I - \rho(\phi)|$ and strictly less complex ($\gamma(\phi') < \gamma(\phi)$). A policy is *kludged* if it is neither L-pure nor R-pure. (See Figure 2.) Note that there is a unique policy ϕ_I^* that is both unkludged and ideal for I; this policy is the I-pure policy with length $\hat{\rho}_I$.

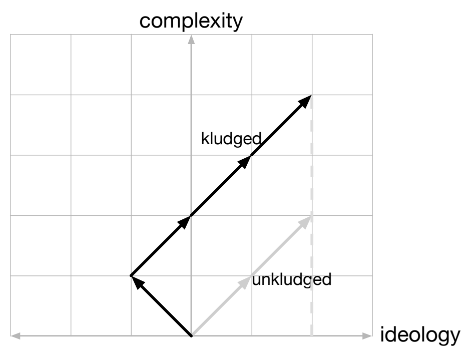

Figure 2: Kludged Policy

2.1 Discussion of the Model

Before proceeding, let us discuss the motivation behind some of our modelling choices.

The assumption that the policymaker may only remove the most recently added rule reflects the premise that there are complementarities between rules: newer rules build upon older rules and rely crucially on the context provided by these older rules, so that removing older rules would render the newer rules incoherent, making the implementation of policy ambiguous and confusing. We make the extreme assumption that such confusion incurs effectively infinite costs, so that any older rule cannot be removed without first removing all newer rules. We conjecture that relaxing this assumption, so that older rules can be removed before newer rules but at a cost, will not change the main insights of the model.

Our model assumes that policymaking is incremental: rules may only be added or removed one at a time. As Hitchins (2008) and Teles (2013) point out, political constraints such as resistance by interest groups (see, e.g., Morris and Coate (1999)) force policymakers to focus their efforts on incremental changes rather than complete overhauls.[1] Building on this interpretation, think of delays in adding or removing rules as being due to resistance from political interest groups (who support or oppose those rules) that has to be overcome before the changes are implemented.

The assumption that $p > q$ captures the premise that there is hysteresis in policymaking, so that rules are easier to add than to remove. This assumption matters for our results: it ensures that player prefer (at least sometimes) to add rules in their favoured direction, rather than remove unfavourable rules. In our setting, there is a natural motivation for this premise: backward dependence implies that only the existing (most recent) rule may be removed, whereas when adding a new rule, there may be multiple potential rules for the policymaker to choose from. This means that the policymaker faces fewer constraints when adding rules than when removing them. Our model captures this point in reduced form, by assuming that the policymaker can add a new rule more quickly than he can remove the most recent rule. Besides backward dependence, other reasons for hysteresis have been extensively discussed and motivated in the literature; for example, Morris and Coate (1999) argue that policies may be easier to enact than remove because, once enacted, interest groups may make policy-specific investments and subsequently fight harder against the removal of these policies.[2]

In the model, the ease with which policy can be modified (as represented by the arrival rates p and q) is independent of the current policy position and of the policymaker's political stance. This assumption is made for tractibility; richer models that take into account the political feasibility of potential changes may yield additional insights, although we expect the main results of the current model to be preserved.

We model the two players' preferences as being diametrically opposed, in the sense that (at least at the origin) a rule that is good for L is bad for R, and vice versa. This assumption is made for parsimony, and in fact makes kludge more difficult to produce in the model: it maximizes each player's motivation to undo rules introduced by his opponent rather than add rules of his own. Accordingly, we expect models with richer player preferences (and a richer space of policies) to preserve our main insights.

2.2 Technical Preliminaries

In this game, the relevant state variable is the combination (ϕ_t, I_t) of policy and the identity of the player in control. With this in mind, we restrict attention to pure-strategy Markov-perfect equilibria whereby for each policy ϕ, each player either selects a single adjacent policy or stagnates at ϕ.

Lemma 1 *A pure-strategy Markov-perfect equilibrium exists.*

[1] Besides political constraints, cognitive limitations may introduce uncertainty about the impact of large-scale policy changes and thus force policymakers to focus on making small 'local' changes to policy (see, e.g., Callander (2011)).

[2] For more in this vein, see Alesina and Drazen (1991)

Each player I's strategy defines a directed graph on the set of policies, whereby $\phi \to \phi'$ whenever I's target at ϕ is ϕ'. We will restrict attention to equilibria where each player's graph is acyclic; this means that each player never returns to a policy that he previously moved away from. This restriction does not have any substantive impact, and eliminates only knife-edge equilibria where players are indifferent between adjacent policies, while simplifying the exposition substantially.

Lemma 2 *There exists a pure-strategy Markov-perfect equilibrium where both players' graphs are acyclic.*

Further, given I's strategy and a policy $\phi_{(0)}$, we define I's *trajectory* $\Phi_I(\phi_{(0)})$ to be the (possibly infinite) sequence of policies $\{\phi_{(0)}, \phi_{(1)}, ..., \phi_{(n)}\}$ such that for each $k \geq 0$, I targets $\phi_{(k+1)}$ when at $\phi_{(k)}$. In other words, $\Phi_I(\phi_{(0)})$ is the sequence of policies (starting from $\phi_{(0)}$) that I will move along on the equilibrium path while he is in control. (For player L, the sequence may be interrupted if he loses control to R before reaching the last policy in his trajectory.)

3 One-Player Game

In this section, we first analyze the subgame for the second player R. This analysis is a useful starting point: it allows us to study optimal policymaking in the absence of strategic interactions between players, and build some intuition for the rest of the analysis.

We start with some notation, followed by basic observations. First, suppose that R's trajectory starting from $\phi_{(0)}$ is $\Phi = \{\phi_{(0)}, \phi_{(1)}, \phi_{(2)}, ..., \phi_{(n)}\}$. While at $\phi_{(k)}$, R will jump to $\phi_{(k+1)}$ with arrival rate p (respectively q) if $\phi_{(k+1)}$ is an extension (respectively truncation) of $\phi_{(k)}$. Denote this arrival rate by $\psi_{(k)}$. For $k < n$ we can write the "asset equation" for R's value function as $(r + \psi_{(k)})V_R(\phi_{(k)}, R) = \pi_R(\phi_{(k)}) + \psi_{(k)}V_R(\phi_{(k+1)}, R)$; or equivalently,

$$V_R\left(\phi_{(k)}, R\right) = \frac{\pi_R(\phi_{(k)}) + \psi_{(k)}V_R(\phi_{(k+1)}, R)}{r + \psi_{(k)}}. \quad (2)$$

Iteratively expanding this expression, we get (defining $\psi_{(n)} = 0$):

$$V_R\left(\phi_{(k)}, R\right) = \sum_{j=k}^{n} \frac{\prod_{i=k}^{j-1} \psi_{(i)}}{\prod_{i=k}^{j}\left(r + \psi_{(i)}\right)} \pi_R\left(\phi_{(j)}\right). \quad (3)$$

In other words, R's value function at $\phi_{(k)}$ is the discounted weighted average of his instantaneous payoffs at each of the policies on his trajectory, starting from $\phi_{(k)}$.

One implication of (3) is that if R extends in equilibrium, he always does so in his favoured direction (so, R only extends at non-ideal policies). Adding a rule in his favoured direction, by moving him towards his ideal, increases his instantaneous payoff; further, it moves him closer to policies that are even closer to his ideal and thus even more lucrative. This observation, and the fact that we focus on acyclic strategies, pins down the form of R's optimal strategy:

Lemma 3 *Fix policy ϕ. Then for some $k \geq 0$ and $k' \geq 0$, R's equilibrium trajectory starting from ϕ is as follows: R will remove k rules from ϕ, then add k' rules in his favoured direction until his ideal is attained, and stagnate thereafter.*

The case where ϕ_t is not R-pure highlights the tradeoff that R faces between attaining his ideal and reducing kludge. The following proposition states that a zealous player will extend (unless his ideal has been reached), whereas a moderate player will undo any existing kludge.

Proposition 1 *Fix all parameters except for ζ_R. Suppose that the current policy ϕ is not R-pure. Then there exists $\underline{\zeta}_R(\phi) > 1$ such that, at ϕ, (i) R extends in his favoured direction if $\zeta_R > \underline{\zeta}_R$; whereas (ii) R undoes if $\zeta_R < \underline{\zeta}_R$.*

Underlying Proposition 1 is the following trade-off. Extending is the fastest way for the player to move in his favoured direction. In comparison, undoing slows the player's progress towards his ideal, but reduces policy complexity. Thus an extremist (who cares greatly about ideological bias relative to complexity) prefers to extend, whereas a moderate (who cares more about complexity relative to bias) prefers to undo. The following example clarifies this intuition in a simplified setting.

Example 1 *Suppose that the starting policy $\phi = \{-1\}$, and that $\hat{\rho}_R = 1$ (i.e., R's ideal is one step to the right of the origin). Then there are two candidates for R's optimal trajectory:*

$$\Phi = \{\{-1\}, \{-1, 1\}, \{-1, 1, 1\}\},$$

where R extends from ϕ until he achieves his ideal, and

$$\Phi' = \{\{-1\}, \{\}, \{1\}\},$$

where R removes the only rule from ϕ to return to the origin, then extends rightward to reach his ideal. Applying (3), the value functions for R given each trajectory are:

$$\begin{aligned} V(\phi, R; \Phi) =& w_{(0)}\pi_R(\phi_{(0)}) + w_{(1)}\pi_R(\phi_{(1)}) + w_{(2)}\pi_R(\phi_{(2)}) \text{ where} \\ w_{(0)} =& \frac{1}{r+p}, w_{(1)} = \frac{p}{(r+p)^2}, w_{(2)} = \frac{p^2}{r(r+p)^2}, \\ V(\phi, R; \Phi') =& w'_{(0)}\pi_R(\phi'_{(0)}) + w'_{(1)}\pi_R(\phi'_{(1)}) + w'_{(2)}\pi_R(\phi'_{(2)}) \text{ where} \\ w'_{(0)} =& \frac{1}{r+q}, w'_{(1)} = \frac{q}{(r+p)(r+q)}, w'_{(2)} = \frac{qp}{r(r+p)(r+q)}. \end{aligned}$$

To compare the two continuation values, note the following facts.

1. *$\pi_R(\phi_{(k)})$ and $\pi_R(\phi'_{(k)})$ are increasing in k: along each trajectory, R's instantaneous payoff increases with each step that he takes.*
2. *for each $k \geq 1$, $\pi_R(\phi'_{(k)}) - \pi_R(\phi_{(k)}) = 2$: the k-th policy in Φ' has the same bias as, but lower complexity than, the corresponding policy in Φ, and thus is more lucrative.*

3. *$w_{(0)} < w'_{(0)}$, whereas for $k \in \{1,2\}$, $w_{(k)} > w'_{(k)}$: compared to the Φ-value function, the Φ'-value function puts more weight on the (less lucrative) early policy ϕ, and less weight on the (more lucrative) later policies. This is because removing rules is slower than adding rules ($q < p$), so R is stuck for a longer time at ϕ under Φ' than under Φ.*

To summarize – the advantage of extending (Φ) over undoing (Φ') is that it allows R to move along the trajectory towards more lucrative policies more quickly; this advantage is increasing in ζ_R. The disadvantage is that each policy after ϕ in Φ is more complex and thus less lucrative than the corresponding policy in Φ'. Consequently, extending is optimal for R if ζ_R is high, whereas undoing is optimal if ζ_R is low.

4 Kludge

In the next few sections, we use the two-player game to study the outcome of political conflict between policymakers. This section shows how conflict between political opponents leads to the emergence and persistence of kludge. We start with the main result. Define $\bar{\zeta}_R = 1 + \frac{2q}{(p-q)\left(\frac{r}{p+r}\right)}$ and $\underline{\zeta}_L = 1 + \frac{2p^2\lambda}{r((p+r)^2+\lambda(2p+r))}$.

Proposition 2 *Suppose both players are sufficiently zealous: $\zeta_L > \underline{\zeta}_L$ and $\zeta_R > \bar{\zeta}_R$. Then*

- *From any starting policy, R's trajectory consists only of right-sided rules, attains R's ideal, and ends there.*
- *Starting from the origin, L's trajectory consists only of left-sided rules and attains (possibly overshooting) L's ideal.*

Consequently, with positive probability, the long-run policy $\lim_{t\to\infty} \phi_t$ is kludged.

When both players are zealous, they add rules in opposite directions. These conflicting rules cancel out in terms of policy bias, but add up in terms of complexity. The result is kludge.

To understand this result, start by remembering (from Proposition 3) that a zealous player R prefers to extend by adding right-sided rules rather than undoing any rules that L put in place. In fact, when R is sufficiently zealous ($\zeta_R > \bar{\zeta}_R$), he will extend towards his ideal at every non-ideal policy.

Such single-minded behavior by R simplifies dramatically the strategic considerations for player L. In this case, L's decision boils down to choosing between stagnating versus adding left-sided rules. Loosely speaking, from L's perspective, neutral and right-sided rules are detrimental: their only effect is to increase complexity, while (weakly) damaging L's ideological position. (In the next section, we'll see how this logic changes if the opponent R is not zealous.) Consequently, L's choice is effectively between adding left-sided rules and stagnating.

Consider the tradeoff involved in adding left-sided rules (relative to stagnating at the origin) from L's perspective. In the short-run, whilst he is in control, L moves policy closer to his ideal. However, in the long-run, R will extend policy rightward to the R-ideal, without first undoing the left-sided rules added by L. In other words, adding L-sided rules adds long-run complexity without improving the long-run position. With this tradeoff in mind, a zealous L thus chooses to add left-sided rules rather than stagnate because he puts greater weight on ideological position over policy complexity.

To highlight the role of political conflict in producing kludge, consider the outcome of the one-player game where R is in control starting from $t = 0$.

Proposition 3 *On the equilibrium path of the one-player game, starting from the origin, R adds rules in his favoured direction (i.e., to the right) until he attains the R-ideal policy, then stagnates there. Consequently, every policy on the equilibrium trajectory is unkludged.*

Proposition 3 describes the outcome in a setting where political competition is absent. Here, R only extends policy in one direction, so kludge does not emerge either in the short or long run. The point is that political conflict plays a crucial role in producing kludge. Specifically, kludge arises when conflicting players add conflicting rules to policy without undoing the rules implemented by their opponents.

Because R never undoes the rules that L previously added, any resulting kludge is persistent. This can be inferred immediately from the observation that player strategies are acyclic.[3] Interestingly, backward-dependence exacerbates the persistence of kludge. As R adds more rules to existing policy, it becomes prohibitively costly for him to remove those rules introduced by L; to do so would require (due to backward depdendence) that he first remove his own rules, and in doing so revert to unfavourable positions that he is unwilling to tolerate. In contrast, in a model without backward dependence, R may undo kludge in the long-run by removing rules added by L even after R has added his own rules.

4.1 Avoiding Kludge: Appeasement

Proposition 9 specifies that kludge emerges if both players are sufficiently zealous. We can produce a partial converse: if either player is sufficiently moderate, then long-run policy is unkludged. Thus, kludges emerge and persist if and (loosely speaking) only if both players are zealous.[4] Define $\underline{\zeta}_R = 1 + \frac{q}{p\left(1-\left(\frac{p}{p+r}\right)^{\hat{\rho}_L+\hat{\rho}_R}\right)}$.

Proposition 4 *If either player is sufficiently moderate ($\zeta_L < \underline{\zeta}_L$ or $\zeta_R < \underline{\zeta}_R$), then the long-run policy $\lim_{t\to\infty} \phi_t$ is unkludged.*

A moderate R will avoid long-run kludge by undoing the rules added by L on the equilibrium path. The case where L is moderate is more interesting. In fact, when

[3] Note that acyclicity is not merely an assumption, and in fact is the generic equilibrium outcome of the model.

[4] Our results focus on the cases where player R has extreme preferences (either high or low ζ_R). In each of these cases, R's optimal strategy has a simple characterization; which in turn simplifies the logic of strategic interaction between the players. The case of intermediate preference intensities is more subtle (and, we think, less interesting), and we defer it for now.

L is moderate and R is zealous, strategic considerations cause L to engage in we call *appeasement*: at the origin, L will stagnate rather than extend leftward. He does so to avoid the production of kludge by R.

Proposition 5 *Suppose L is sufficiently moderate and R is sufficiently zealous: $\zeta_L < \underline{\zeta}_L$ and $\zeta_R > \bar{\zeta}_R$. Then L will stagnate at the origin, and the long-run policy* $\lim_{t\to\infty} \phi_t$ *is unkludged.*

Although L would increase his instantaneous payoff by adding left-leaning rules starting from the origin (and in fact would do so in the absence of political competition), the prospect of being succeeded by a zealot induces him to avoid adding any rules at all. L anticipates that R, upon ascending to power, will not undo any existing rules but instead will add right-leaning rules until he reaches his ideal. Thus, by adding rules while he is in control, L shifts the policy position in his favour in the short-run, but increases the amount of long-run kludge. A sufficiently moderate L (who cares little about policy position relative to policy complexity) thus chooses not to add any rules at all, effectively conceding the policymaking process entirely to his opponent.

5 Obstructionism

Intentional complexity – in the form of designs or rules that are intentionally made to be excessively complex or confusing – appears in fields ranging from patent law to software development. Meanwhile, a number of literatures have discussed the phenomenon of obstructionism, whereby agents take actions that does not improve their current payoffs from policy, but instead make it more difficult for their opponents to make progress. In this section, we show how obstructionism manifests in public policy as a form of intentional complexity, in the sense that policymakers deliberately implement excessively complex policies to obstruct their opponents.

Two observations before diving in. First, in our model, only the first player L engages in obstructionism, if at all: the second player R has no reason to engage in strategic behavior of any sort. Second, only zealous players will engage in obstructionism: because zealous players prioritize ideological position over policy complexity, they are willing to tolerate the increased complexity that arises from obstructionist behavior. So, we focus on the case where player L is zealous.

5.1 Intentional Complexity

In this section, we show that obstructionism against a moderate opponent optimally takes a form we term *intentional complexity.* Specifically, the following proposition shows that a zealous first player L will, after reaching his ideal policy, add neutral rules that increase policy complexity without improving the policy's ideological position.

Proposition 6 *Fix all parameters except ζ_L and ζ_R. Then there exists $\bar{\zeta}_L$ and $\underline{\zeta}_R$ such that if L is a zealot ($\zeta_L > \bar{\zeta}_L$) and R is a moderate ($\zeta_R < \underline{\zeta}_R$), then along L's policy* *sequence, he extends leftward until he reaches the L-pure, L-ideal policy, then (for one or more steps) extends upward by adding neutral rules.*

At first glance, this result may seem surprising: by adding neutral rules, L is moving to a more complex policy without any improvement in his immediate ideological position. To understand why L adds these neutral rules, remember that R is moderate, and thus will attempt to undo any rules that L had added. Thus L can delay R's rightward movement simply by adding more rules that R is compelled to undo. In other words, these ostensibly pointless rules that L adds serve as a bulwark against the future advance of R. And the least costly way for L to construct this bulwark is to add neutral rules after attaining his ideal policy (otherwise, to extend in either direction from his ideal policy would entail a deterioration in his ideological position).

5.2 Strategic Extremism

When facing a zealous player R, player L may engage in a different form of obstructionism that we term *strategic extremism*. Specifically, L may extend leftward even after attaining his ideal policy, resulting in 'extremist' policies that lie left of L's ideal.

Proposition 7 *Suppose that the first player L is vulnerable (i.e, $\lambda > 2p$), and both players are patient. If both players are zealots, then L will extend leftward at L's ideal policy. That is, there exists an $\underline{r} > 0$ such that for all $r < \underline{r}$, there exist $\bar{\zeta}_L$ and $\bar{\zeta}_R$ and if $\zeta_L > \bar{\zeta}_L$ and $\zeta_R > \bar{\zeta}_R$, then L will extend leftward at L's ideal policy.*

Proposition 7 tells us that when facing a zealous opponent, L 'overshoots' his ideal by adding left-biased rules. This differs from how L behaves when he faces a moderate opponent, in which case he doesn't overshoot but instead adds neutral rules while staying at his ideal.

With both types of obstructionism, L is attempting to delay R's progress towards his ideal. However, the way that L obstructs R depends on whether R is a moderate (Proposition 6) or a zealot (Proposition 7). Recall that when R is a moderate, L can slow his progress by adding neutral rules that R is compelled to remove before starting to move rightward. Such a strategy fails when R is a zealot, because R will extend rightward rather than undoing existing rules – so neutral rules do nothing to impede R's progress. Instead, L can delay R by adding left-biased rules, so that R's starting point (when he takes control) is further left. Doing so ensures that R takes a longer time to reach ideological positions that are relatively unfavourable to L. Thus, by engaging in strategic extremism, L profitably delays the reduction in his own payoffs that occurs as R moves rightward.

Notice that Proposition 7 applies only when player L is neither too patient nor too impatient. Why? If L engages in strategic extremism, then compared to a strategy whereby he stagnates at his ideal policy, he is:

- worse off in the short run, while the policy remains left of his ideological position, before R takes control and extends rightward.

- better off in the medium run, when the policy lies between L's ideological position and R's ideological position, because overshooting delays R's rightward progress;
- worse off in the long run, when R has attained his ideological bliss point, because by extending leftward, L introduces more complexity into the policy over the long run.

Consequently, the medium-run advantages of strategic extremism overwhelm the short- and long-run drawbacks only if L is neither too patient nor too impatient.

Propositions 6 and 7 highlight the point that the optimal form of obstructionism depends on your opponent's preferences; intentional compleixty is optimal against a moderate opponent while strategic extremism is optimal against a zealous opponent.

6 Frictions

Now, we turn to consider the effect of frictions in policymaking that constrain the ability of policymakers to add or remove rules to policy. For example, U.S. policymakers face a high-friction environment. In the U.S. political system, there are a multitude of veto points in the legislative process, and it is difficult for the party in power to successfully shepherd legislative proposals through these veto points; with the result that attempts to pass or undo legislation take longer to succeed. These frictions are often mooted as a positive feature of U.S. democracy: by making it difficult to change policy, shifts in policy position may be avoided, thus reducing policy bias.

The following proposition points out that such frictions come at a cost: high-friction political systems may induce players to prioritize adding over removing rules, and thus may result in the emergence and persistence of kludge. Conversely, low-friction political systems avoid kludge because they ensure that (i) players do not engage in obstructionary behavior, and that (ii) players undo undesirable rules, so that long-run kludge is avoided.

Proposition 8 *Let $p = \hat{p}/\chi$ and $q = \hat{q}/\chi$. Consider the two-player game, and fix all parameters except the degree of friction χ. If R is sufficiently moderate ($\zeta_R < \frac{\hat{p}+\hat{q}}{\hat{p}-\hat{q}}$), then long-run policy is always unkludged for all χ. On the other hand, if $\zeta_R > \frac{\hat{p}+\hat{q}}{\hat{p}-\hat{q}}$, then*

- *There exists $\underline{\chi} > 0$ such that for $\chi < \underline{\chi}$, on the equilibrium path, the first player L extends in direction d_L until he reaches the L-ideal policy, at which he stagnates; whereas the second player R undoes any rules that L added, then extends in direction d_R until he reaches the R-ideal policy, at which he stagnates. Consequently, every policy on the equilibrium path is unkludged.*
- *There exists $\overline{\chi} > 0$ such that for $\chi > \overline{\chi}$, each player I extends in his favoured direction at every policy that does not attain his ideal. Consequently, the long-run policy is kludged with positive probability.*

The intuition underlying this result is as follows. When frictions are high, changes in policy are relatively slow in arriving. The policymaker anticipates that it is unlikely for multiple policy changes to occur within his relevant time horizon, and thus focuses on maximizing the payoff from the next policy change that he is attempting to make.

Given our assumption that the second policymaker R not too moderate, each of them is better off extending rather than to undoing (so as to get to a more favourable ideological position more quickly). Kludge thus arises with positive probability because the second player R will not attempt to undo any rules added by the first player.

On the other hand, when frictions are low, changes in policy arrive rapidly. Each policymaker X anticipates that he is likely to achieve any sequence of policy changes he wishes to make in a negligibly short period of time; thus, he chooses the sequence that arrives at his ideal policy (i.e. undoing any existing rules, then extending in his favoured direction). Further, the first player L avoids an obstructionary strategy because he anticipates that any obstructionary rules that he adds will be dismantled by R almost immediately, and thus that such a strategy would be counterproductive (because it reduces L's payoffs while he is in control).

A caveat: although Proposition 8 emphasizes that kludge emerges and persists only in sufficiently high-friction systems, the impact of friction on the expected amount of long-run kludge,

$$E\left[\lim_{t\to\infty}(\gamma(\phi_t) - |\psi(\phi_t)|)\right],$$

is nonmonotone. In fact, as the degree of friction γ goes to infinity, the probability that the long-run policy is kludged goes to zero. This is because increased friction reduces the ability of players to change policy; in particular, in a high friction envronment, L is unlikely to successfully add (left-biased) rules before R gains control, so policy is unlikely to become kludged. That said, arbitrarily high frictions are unrealistic in practice, because of the need for policy to be adapted readily to shocks in the political environment (e.g. economic, cultural or technological changes).[5] So a more careful interpretation of Proposition 8 is that, at least within an intermediate range, an increase in policymaking friction may induce an increase in policy kludge.

7 Conclusion

In this paper, we have worked out a simple model of sequential, path-dependent policymaking. The key assumption is technological: when undoing existing policy, newer rules have to be removed before older rules. The analysis focuses on the effect of political conflict between policymakers. We show that a number of interesting phenomena arise from strategic interactions in the model: how kludge emerges when zealots conflict, various forms of obstructionary behavior, and the impact of frictions on long-run policy outcomes.

Throughout the paper, we have emphasized the application of our model to public policy. However, we believe that our model may also be relevant to other settings, such as the politics of organizational policy-making. The insights we derive in the model can be straightforwardly reinterpreted for an organizational context; for example, our results on long-run kludge suggest that political conflict between different factions within an

[5] To capture this point, a richer model could incorporate random shocks to either preferences or payoffs that require policy to adapt in response.

organization may give rise to inefficiently bureaucratic routines and procedures within the organization.[6]

Returning to the topic of public policy, how might the insights of this model be applied to the design of political institutions? In particular, what is the optimal degree of political competition? Interpret the number of players in our model in terms of the degree of political competition, the one-player game corresponds to an uncompetitive setting (e.g. an autocracy) whereas the two-player game corresponds to a competitive setting (e.g. a democracy). This interpretation suggests that an increase in political competition may reduce the occurrence of extremely biased policies, but increased competition comes at the cost of increased long-run policy complexity.[7] Thus, a farsighted social planner who abhors policy complexity may prefer an autocratic political system over a democracy. That said, that the detrimental effects of political competition on complexity are moderated when policymakers are not too zealous. This suggests that in culturally homogenous societies, where competing political parties have moderate policy preferences, it may be optimal to implemental democratic political institutions. On the other hand, culturally fragmented or heterogenous societies may be better off with autocratic institutions, so as to avoid the emergence of policy kludge.

8 Appendix

Some notation: $V_I(\phi, J)$ is the equilibrium continuation value for player I when the current policy is ϕ and player J is in control. $V_I(\phi, I; \Phi)$ is the equilibrium continuation value for player I, conditional on pursuing trajectory Φ and and on the current policy ϕ lying on the trajectory Φ, when he is in control.

Proof of Lemma 1 Lippman (1976), Theorem 7; notice that we apply the more general assumptions from Lippman (1975). These more general assumptions do not affect the argument. ■

Proof of Lemma 2 First, we'll show that in any equilibrium, R's strategy is cyclic. To show that R's strategy is acyclic, assume otherwise and consider an policy ϕ_τ, $\tau \in (t, t')$, that is adjacent to ϕ_t. Moreover, our restriction to pure strategies implies that for any τ', $\phi_{\tau'}$ is either ϕ_t or ϕ_τ for all $\tau' \in (t, t')$. Since $\eta_R > 1$, adjacent policies produce different instantaneous payoffs. That is, $\pi_R(\phi_t) \neq \pi_R(\phi_\tau)$. Therefore, R is strictly better off stagnating at ϕ_t if $\pi_R(\phi_t) > \pi_R(\phi_\tau)$, and at ϕ_τ if $\pi_R(\phi_t) < \pi_R(\phi_\tau)$ than moving to $\phi_{t'}$, a contradiction.

Now, fix any pure-strategy equilibrium, and consider L's strategy. If L undoes from policy ϕ to policy ϕ', then it must be that L switches indefinitely between ϕ and ϕ',

[6] Relatedly, one might expect to see more kludge in the code of teams of software developers where members of the team have conflicting ideals.

[7] Note that an increase in policymaking frictions involves a similar tradeoff: increased friction results in slower movement towards extreme policies, but makes policymakers more inclined to implement kludged policies.

and that L is indifferent between ϕ and ϕ'. But then L would receive the exact same continuation value from stagnating at ϕ' instead (and also, the continuation values starting from any other policy would remain unchanged as well). Thus the alternative pure strategy where L stagnates at ϕ is also an equilibrium. ■

Proof of Proposition 1 Fix ϕ. We start with the case where ϕ is neither R-ideal nor R-pure. Let σ_k be the strategy of removing k rules from ϕ and then add R-favoured rules until R's ideal is reached, and $(\phi^k_{(0)}, \phi^k_{(1)}, ..., \phi^k_{(n_k)})$ be the associated trajectory of $n_k + 1$ steps, where $\phi^k_{(0)} = \phi$. Lemma 3 ensures that we can restrict attention to strategies σ_k for $k \in [0, \gamma(\phi)]$ such that $\pi\left(\phi^k_{(k)}\right) < \pi\left(\phi^k_{(k+1)}\right)$. Let $V_k\left(\tilde{\phi}\right)$ be the value function of taking strategy σ_k at an on-trajectory policy $\tilde{\phi}$. First we argue that we are done if we show that (a) $V_k(\phi)$ is linear in ζ_R, (b) $\lim_{\zeta_R \to 1} V_k(\phi) < 0$ and $\lim_{\zeta_R \to \infty} V_k(\phi) = \infty$. This is because (a) and (b) together imply that for any k, there exists $\underline{\zeta}^k_R(\phi) \in (1, \infty)$ such that $V_0(\phi) - V_k(\phi) < 0$ if and only if $\zeta_R < \underline{\zeta}^k_R(\phi)$. Therefore, if we define $\underline{\zeta}_R(\phi) = \min_{k \in [0, \gamma(\phi)]} \left\{\underline{\zeta}^k_R(\phi)\right\}$, then we can conclude that R will optimally play σ_0 (i.e. extend at ϕ) if and only if $\zeta_R > \underline{\zeta}_R(\phi)$; otherwise he will prefer some σ_k, and undo at ϕ instead.

Now we prove (a). Note that the value function V_k, given strategy σ_k, for an on-trajectory policy satisfies the equation

$$rV_k\left(\phi^k_{(m)}\right) = \pi_R\left(\phi^k_{(m)}\right) + \psi^k_{(m)}\left(V_k\left(\phi^k_{(m+1)}\right) - V_k\left(\phi^k_{(m)}\right)\right),$$

where $\psi^k_{(m)} = p$ if $\phi^k_{(m+1)}$ is extended from $\phi^k_{(m)}$, $\psi^k_{(m)} = q$ if $\phi^k_{(m+1)}$ is undone from $\phi^k_{(m)}$, and $\psi^k_{(m)} = 0$ if $m = n_k$. This can be rewritten as

$$V_k\left(\phi^k_{(m)}\right) = \frac{\pi_R\left(\phi^k_{(m)}\right) + \psi^k_{(m)} V_k\left(\phi^k_{(m+1)}\right)}{r + \psi^k_{(m)}}; \tag{4}$$

further, decomposing R's instantaneous payoff for each policy into the "position" component and the "complexity" component:

$$\pi_R(\phi) = \zeta_R \pi^\rho_R(\phi) + \pi^\gamma_R(\phi) = -\zeta_R \left|\hat{\rho}_R - \rho(\phi)\right| - \gamma(\phi),$$

we can decompose $V_k(\phi)$ into $\zeta_R V^\rho_k(\phi) + V^\gamma_k(\phi)$ where

$$V^\rho_k\left(\phi^k_{(m)}\right) = \frac{\pi^\rho_R\left(\phi^k_{(m)}\right) + \psi^k_{(m)} V^\rho_k\left(\phi^k_{(m+1)}\right)}{r + \psi^k_{(m)}}, V^\gamma_k\left(\phi^k_{(m)}\right) = \frac{\pi^\gamma_R\left(\phi^k_{(m)}\right) + \psi^k_{(m)} V^\gamma_k\left(\phi^k_{(m+1)}\right)}{r + \psi^k_{(m)}}.$$

This proves that $V_k(\phi)$ is linear in ζ_R.

Next, we prove (b). We start with $\lim_{\zeta_R \to 1} V_0(\phi) - V_k(\phi) < 0$. To see this, note that along the trajectory corresponding to σ_0, i.e.

$$\lim_{\zeta_R \to 1} \pi^\rho_R\left(\phi^0_{(m)}\right) + \pi^\gamma_R\left(\phi^0_{(m)}\right) = \lim_{\zeta_R \to 1} \pi^\rho_R\left(\phi^0_{(0)}\right) + \pi^\gamma_R\left(\phi^0_{(0)}\right) = \pi_R(\phi), \text{ for all } m. \tag{5}$$

That is, $\pi_R^\rho\left(\phi_{(m)}^0\right)+\pi_R^\gamma\left(\phi_{(m)}^0\right)$ is constant when $\zeta_R=1$. Therefore, $\lim_{\zeta_R\to 1} V_0(\phi)=\frac{\pi_R(\phi)}{r}$. On the other hand, since ϕ is not pure $\lim_{\zeta_R\to 1}\pi_R^\rho\left(\phi_{(m)}^k\right)+\pi_R^\gamma\left(\phi_{(m)}^k\right)\geq \pi_R^\rho(\phi)+\pi_R^\gamma(\phi)+1$ for some m. Moreover, since $\lim_{\zeta_R\to 1}\pi_R^\rho\left(\phi_{(m)}^k\right)+\pi_R^\gamma\left(\phi_{(m)}^k\right)$ is weakly increasing in m, $\lim_{\zeta_R\to 1} V_k(\phi)>\frac{\pi_R(\phi)}{r}$. Therefore, our claim that $\lim_{\zeta_R\to 1} V_0(\phi)-V_k(\phi)<0$ follows.

We show that $\lim_{\zeta_R\to\infty} V_0(\phi)-V_k(\phi)=\infty$. Note that this follows when $V_0^\rho\left(\phi_{(0)}^0\right)=V_k^\rho(\phi)>V_0^\rho(\phi)=V_0^\rho\left(\phi_{(0)}^k\right)$. Now we claim that $V_0^\rho\left(\phi_{(m)}^0\right)>V_k^\rho\left(\phi_{(m)}^k\right)$ for all m by induction. For $m=n_0$, note that $\pi_R^\rho\left(\phi_{n_0}^0\right)=\pi_R^\rho\left(\phi_{n_k}^k\right)$, and $\pi_R^\rho\left(\phi_{n_k}^k\right)>\pi_R^\rho\left(\phi_m^k\right)$ for all $m\in[n_0,n_k]$. Therefore, $V_0^\rho\left(\phi_{(n_0)}^0\right)=V_k^\rho\left(\phi_{(n_k)}^k\right)>V_k^\rho\left(\phi_{(n_0)}^k\right)$. Next, suppose that $V_0^\rho\left(\phi_{(m)}^0\right)>V_k^\rho\left(\phi_{(m)}^k\right)$ for all $m>\bar{m}+1$ for some $\bar{m}$. Note that

$$
\begin{aligned}
&V_0^\rho\left(\phi_{(\bar{m})}^0\right)\\
&=\frac{r}{r+\psi_{(\bar{m})}^0}\frac{\pi_R\left(\phi_{(\bar{m})}^0\right)}{r}+\frac{\psi_{(\bar{m})}^0}{r+\psi_{(m)}^0}V_0^\rho\left(\phi_{(\bar{m}+1)}^0\right)\\
&=\frac{r}{r+\psi_{(\bar{m})}^0}\frac{\pi_R\left(\phi_{(\bar{m})}^0\right)}{r}+\left(1-\frac{r}{r+\psi_{(\bar{m})}^0}\right)V_0^\rho\left(\phi_{(\bar{m}+1)}^0\right)\\
&=\frac{r}{r+\psi_{(\bar{m})}^k}\frac{\pi_R\left(\phi_{(\bar{m})}^0\right)}{r}+\left(\frac{r}{r+\psi_{(\bar{m})}^k}-\frac{r}{r+\psi_{(\bar{m})}^0}\right)\left(V_0^\rho\left(\phi_{(\bar{m}+1)}^0\right)-\frac{\pi_R\left(\phi_{(\bar{m})}^0\right)}{r}\right)+\frac{\psi_{(\bar{m})}^k}{r+\psi_{(m)}^k}V_0^\rho\left(\phi_{(\bar{m}+1)}^0\right)
\end{aligned}
$$

Since $\psi_{(m)}^0\equiv p\geq\psi_{(m)}^k$ for all m, $\frac{r}{r+\psi_{(\bar{m})}^k}>\frac{r}{r+\psi_{(\bar{m})}^0}$. Also, note that $V_0^\rho\left(\phi_{(\bar{m}+1)}^0\right)>\frac{\pi_R\left(\phi_{(\bar{m})}^0\right)}{r}$ because $\pi_R^\rho\left(\phi_{(m)}^0\right)$ is increasing in m, and $V_0^\rho\left(\phi_{(n_0)}^0\right)=\frac{\pi_R^\rho\left(\phi_{(n_0)}^0\right)}{r}$. Therefore,

$$
\begin{aligned}
&\frac{r}{r+\psi_{(\bar{m})}^k}\frac{\pi_R\left(\phi_{(\bar{m})}^0\right)}{r}+\left(\frac{r}{r+\psi_{(\bar{m})}^k}-\frac{r}{r+\psi_{(\bar{m})}^0}\right)\left(V_0^\rho\left(\phi_{(\bar{m}+1)}^0\right)-\frac{\pi_R\left(\phi_{(\bar{m})}^0\right)}{r}\right)+\frac{\psi_{(\bar{m})}^k}{r+\psi_{(m)}^k}V_0^\rho\left(\phi_{(\bar{m}+1)}^0\right)\\
&\geq\frac{r}{r+\psi_{(\bar{m})}^k}\frac{\pi_R\left(\phi_{(\bar{m})}^0\right)}{r}+\frac{\psi_{(m)}^k}{r+\psi_{(m)}^k}V_0^\rho\left(\phi_{(\bar{m}+1)}^0\right).
\end{aligned}
$$

Then, $\pi_R\left(\phi_{(m)}^0\right)>\pi_R\left(\phi_{(m)}^k\right)$ and the induction hypothesis that $V_0^\rho\left(\phi_{(\bar{m}+1)}^0\right)>$

$V_k^\rho\left(\phi^k_{(\bar{m}+1)}\right)$ together imply

$$\begin{aligned}
&V_0^\rho\left(\phi^0_{(\bar{m})}\right)\\
&\geq \frac{r}{r+\psi^k_{(\bar{m})}}\frac{\pi_R\left(\phi^0_{(\bar{m})}\right)}{r}+\frac{\psi^k_{(m)}}{r+\psi^k_{(m)}}V_0^\rho\left(\phi^0_{(\bar{m}+1)}\right)\\
&> \frac{r}{r+\psi^k_{(\bar{m})}}\frac{\pi_R\left(\phi^k_{(\bar{m})}\right)}{r}+\frac{\psi^k_{(m)}}{r+\psi^k_{(m)}}V_0^\rho\left(\phi^k_{(\bar{m}+1)}\right)\\
&= V_k^\rho\left(\phi^k_{(\bar{m})}\right).
\end{aligned}$$

This proves $V_k^\rho(\phi) > V_0^\rho(\phi)$.

The case where ϕ is neither R-ideal nor R-pure is completely analogous. ■

Lemma 4 *If $\zeta_R > \bar{\zeta}_R$, then R extends in his favoured direction at any policy ϕ that is not R-ideal.*

Proof of Lemma 4 We adopt the terminology of Proposition 1. Choose $\bar{\zeta}_R = 1 + \frac{2q}{(p-q)\left(\frac{r}{p+r}\right)}$. We'll proceed by induction over k on the following statement: If $\zeta_R > \bar{\zeta}_R$, then R extends in his favoured direction at any policy with complexity $\leq k$. Consider a policy $\phi = (r_{(1)}, ..., r_{(k+1)})$ with complexity $\gamma(\phi) = k+1$. By the induction hypothesis, R will optimally extend the truncated policy $\phi' = (r_{(1)}, ..., r_{(k)})$ in his favoured direction; so R will remove at most one rule from ϕ before extending. Thus we may restrict attention to σ_0 (whereby R extends from ϕ immediately) and σ_1 (whereby R removes one rule from ϕ to get ϕ', then extends from ϕ'). In particular, R prefers to extend from ϕ if and only if $V_0(\phi) > V_1(\phi)$. Let $n = |\hat{\rho}_R - \rho(\phi)|$; note that $n \geq 1$. As a preliminary step: applying (4) iteratively, we get

$$V_k(\phi) = \sum_{m=0}^{n_k}\left(\frac{1}{r+\psi^k_{(m)}}\prod_{j=0}^{m-1}\frac{\psi^k_{(j)}}{r+\psi^k_{(j)}}\pi\left(\phi^k_{(m)}\right)\right)$$

where $\psi^k_{(j)} = \begin{cases} p : j \geq k \\ q : j < k \end{cases}$. Note that we only have to consider the cases $r_{(k+1)} = -1$ and $r_{(k+1)} = 0$: σ_1 cannot be optimal for the case $r_{(k+1)} = 1$, as it would produce an cyclic trajectory for R.

$$\begin{aligned}
V_0(\phi) &= \sum_{m=0}^{n_0-1}\left(\frac{1}{r+p}\left(\frac{p}{r+p}\right)^m(-\zeta_R(n_0-m)-(\gamma(\phi)+m))\right)+\frac{1}{r}\left(\frac{p}{r+p}\right)^{n_0}(-(\gamma(\phi)+n_0)),\\
V_1(\phi) &= \frac{1}{r+q}(-\zeta_R n_0-\gamma(\phi))\\
&\quad+\sum_{m=1}^{n_0-1}\left(\frac{1}{r+p}\frac{q}{r+q}\left(\frac{p}{r+p}\right)^{m-1}(-\zeta_R(n_0-m)-(\gamma(\phi)+m-2))\right)\\
&\quad+\frac{1}{r}\frac{q}{r+q}\left(\frac{p}{r+p}\right)^{n_0-1}(-(\gamma(\phi)+n_0-2)).
\end{aligned}$$

Some calculations reveal that

$$V_0(\phi)-V_1(\phi)=\frac{-2q+(\zeta_R-1)(p-q)\left(1-\left(\frac{p}{p+r}\right)^{n_0}\right)}{r(q+r)},$$

so R prefers to extend rather than undo (equivalently, $V_0(\phi)\geq V_1(\phi)$) if $\zeta_R\geq 1+\frac{2q}{(p-q)\left(1-\left(\frac{p}{p+r}\right)^{n_0}\right)}$. Since $\bar{\zeta}_R\geq 1+\frac{2q}{(p-q)\left(1-\left(\frac{p}{p+r}\right)^{n_0}\right)}$, we have the required result for the case $r_{(k+1)}=-1$. The case $r_{(k+1)}=0$ is very similar: we get

$$V_0(\phi)-V_1(\phi)=\frac{-q+(\zeta_R-1)p\left(1-\left(\frac{p}{p+r}\right)^{n_0}\right)}{r(q+r)},$$

so R prefers to extend rather than undo (equivalently, $V_0(\phi)\geq V_1(\phi)$) if $\zeta_R\geq 1+\frac{q}{p\left(1-\left(\frac{p}{p+r}\right)^{n_0}\right)}$. Again, since $\bar{\zeta}_R>1+\frac{q}{p\left(1-\left(\frac{p}{p+r}\right)^{n_0}\right)}$, we have the required result. ■

Lemma 5 *At any policy, player L's trajectory is a finite sequence.*

Proof. Denote L's trajectory, starting from the origin, as $\phi_{(0)},\phi_{(1)},\ldots$. Choose $\tilde{k}=\zeta_R(\hat{\rho}_L+\hat{\rho}_R)+\hat{\rho}_R$. We claim that the length n of L's trajectory does not exceed $\tilde{k}$, which is equivalent to claiming that no policy in L's trajectory has length $\geq\tilde{k}$.

The first step to establishing this claim is to show that $V_L(\phi_{(0)})\geq-\tilde{k}/r$, where $\phi_{(0)}$ is the origin. To see this, note that if L stagnates at the origin, then he ensures (via Lemma 3) that only R-pure policies are attained on the continuation path. Amongst these policies, L's instantaneous payoff is minimized at the R-ideal policy, in which case $\pi_L=-\zeta_R(\hat{\rho}_L+\hat{\rho}_R)-\hat{\rho}_R=-\tilde{k}$; this implies that a lower bound for L's continuation value at the origin is $V_L(\phi_{(0)})>-\tilde{k}/r$.

Choose m such that $\frac{m\zeta_R}{r+p+\lambda}>\tilde{k}/r$. Note that $V_L(\phi_{(m)})=\frac{\pi_L(\phi_{(m)})}{r+p+\lambda}+\frac{p}{r+p+\lambda}V_L(\phi_{(m+1)})+\frac{\lambda}{r+p+\lambda}V_{LR}(\phi_{(m)})<\frac{\pi_L(\phi_{(m)})}{r+p+\lambda}\leq-\tilde{k}/r\leq V_L(\phi_{(0)})$, where $V_{LR}(\phi_{(m)})$ is L's continuation payoff at policy $\phi_{(m)}$ with R in control. But this contradicts the fact that $V_L(\phi_{(n)})$ must be weakly increasing in n (remember that L is only willing to extend from $\phi_{(n)}$ to $\phi_{(n+1)}$ if he increases his continuation value in doing so). We conclude that, by contradiction, L's trajectory is finite. ■

Lemma 6 *If $\zeta_R > \bar{\zeta}_R$, then every policy on L's trajectory, starting from the origin, adds a left-leaning rule to the previous policy; equivalently, every policy on L's trajectory consists solely of left-leaning rules.*

Proof of Lemma 6 Consider, for L, a finite trajectory $\Phi = (\phi_{(0)}, \phi_{(1)}, ..., \phi_{(n)})$ starting from the origin. This trajectory induces a sequence of rules $(d_{(1)}, d_{(2)}, ...d_{(n)})$ such that $\phi_{(k)} \equiv (d_{(1)}, ..., d_{(k)})$. Suppose that there exists at least one neutral or right-biased rule in the sequence. Then unless all rules in the sequence are right-biased (Case 1), there must exist m and $m' > m$ such that $\rho(\phi_{(m')}) = \rho(\phi_{(m)})$ (Case 2). We will show that in each case, there is an alternative trajectory (starting from the origin) that L strictly prefers.

We'll start with Case 2. By definition, $\phi_{(m')}$ is more complex than $\phi_{(m)}$; let $\hat{\gamma} = \gamma(\phi_{(m')}) - \gamma(\phi_{(m)})$. Let's construct an alternative trajectory for L: $\Phi' = (\phi_{(0)}, \phi_{(1)}, ..., \phi_{(m-1)}, \phi_{[m']}, \phi_{[m'+1]}, ..., \phi_{[n]})$ with $\phi_{(m)} = \phi_{[m']}$. This alternative trajectory Φ' is defined by a sequence of rules that is identical to the equilibrium rule sequence for the first m rules and the last $n - m'$ rules, but omits all rules in between: $(d_{(1)}, d_{(2)}, ..., d_{(m)}, d_{(m'+1)}, ..., d_{(n)})$. We claim that under Φ', L's continuation value at any policy $\phi_{[k]}$ (denoted by $V_L^{\Phi'}\left(\phi_{[k]}, L\right)$) with $k \geq m'$ is strictly greater than his continuation value at $\phi_{(k)}$ under Φ ($V_L^{\Phi}(\phi_{(k)}, L)$). This claim, once verified, then implies that $V_L^{\Phi'}(\phi_{(k)}, L) > V_L^{\Phi}(\phi_{(k)}, L)$ for all $k < m$, and thus that L strictly prefers Φ' to Φ.

To keep notation uncluttered, we'll perform the calculation for $k = m'$; the case $k > m'$ is identical. For each $l \geq m'$, let $\left(\phi^l_{[0]}, \phi^l_{[1]}, ..., \phi^l_{[n_l]}\right)$ be R's trajectory starting from $\phi_{[l]} = \phi^l_{[0]}$, and let $\left(\phi^l_{(0)}, \phi^l_{(1)}, ..., \phi^l_{(n_l)}\right)$ be R's trajectory starting from $\phi_{(l)} = \phi^l_{(0)}$. Notice that $\pi_L(\phi^l_{[k]}) \equiv \pi_L(\phi^l_{(k)}) + \hat{\gamma}$, with a special case being $\pi_L(\phi_{[l]}) \equiv \pi_L(\phi_{(l)}) + \hat{\gamma}$. We now show by induction that $V_L(\phi^l_{[k]}, R) \equiv V_L(\phi^l_{(k)}, R) + \frac{\hat{\gamma}}{r}$ for $k \geq 0$. Start with the observation that

$$
\begin{aligned}
V_L(\phi^l_{(n_l)}, R) &= \frac{\pi_L(\phi^l_{(n_l)})}{r}, \\
V_L(\phi^l_{[n_l]}, R) &= \frac{\pi_L(\phi^l_{[n_l]})}{r} = V_L(\phi^l_{(n_l)}, R) + \frac{\hat{\gamma}}{r},
\end{aligned}
$$

so the claim holds for $k = n_l$. Now, suppose that the induction claim holds for $k > 0$. Then the claim holds for $k - 1$ as well:

$$
\begin{aligned}
V_L(\phi^l_{(k-1)}, R) &= \frac{\pi_L(\phi^l_{(k-1)})}{r+p} + \frac{p}{r+p} V_L(\phi^{l-1}_{(k)}, R), \text{ and} \\
V_L(\phi^l_{[k-1]}, R) &= \frac{\pi_L(\phi^l_{[k-1]})}{r+p} + \frac{p}{r+p} V_L(\phi^l_{[k]}, R) \\
&= \frac{\pi_L(\phi^l_{(k-1)}) + \hat{\gamma}}{r+p} + \frac{p}{r+p}\left(V_L(\phi^l_{[k]}, R) + \frac{\hat{\gamma}}{r}\right) \\
&= V_L(\phi^l_{(k-1)}, R) + \frac{\hat{\gamma}}{r}.
\end{aligned}
$$

Our claim thus holds by induction. We then perform a further round of induction to show that $V_L^{\Phi'}(\phi_{[0]}^{m'}, L) \equiv V_L(\phi_{(0)}^{m'}, L) + \frac{\hat{\gamma}}{r}$ for $l \geq m'$. Start with

$$\begin{aligned}
V_L(\phi_{(0)}^n, R) &= \frac{\pi_L(\phi_{(0)}^n)}{r+\lambda} + \frac{\lambda}{r+\lambda} V_L(\phi_{(0)}^n, R), \\
V_L(\phi_{[0]}^n, R) &= \frac{\pi_L(\phi_{[0]}^n)}{r+\lambda} + \frac{\lambda}{r+\lambda} V_L(\phi_{[0]}^n, R) \\
&= \frac{\pi_L(\phi_{(0)}^n) + \hat{\gamma}}{r+\lambda} + \frac{\lambda}{r+\lambda}\left(V_L(\phi_{(0)}^n, R) + \frac{\hat{\gamma}}{r}\right) \\
&= V_L(\phi_{(0)}^n, R) + \frac{\hat{\gamma}}{r}.
\end{aligned}$$

so the claim holds for $l = n$. Now, suppose that the induction claim holds for $l > m'$. Then the claim holds for $l - 1$ as well:

$$\begin{aligned}
V_L^{\Phi}(\phi_{(l-1)}, L) &= \frac{\pi_L(\phi_{(l-1)})}{r+\lambda+p} + \frac{p}{r+\lambda+p} V_L^{\Phi}(\phi_{(l)}) + \frac{\lambda}{r+\lambda+p} V_L(\phi_{(0)}^{l-1}, R), \\
V_L^{\Phi'}(\phi_{[l-1]}) &= \frac{\pi_L(\phi_{[l-1]})}{r+\lambda+p} + \frac{p}{r+\lambda+p} V_L^{\Phi'}(\phi_{[l]}) + \frac{\lambda}{r+\lambda+p} V_L(\phi_{[0]}^{l-1}, R) \\
&= \frac{\pi_L(\phi_{(l-1)}) + \hat{\gamma}}{r+\lambda+p} + \frac{p}{r+\lambda+p}\left(V_L^{\Phi}(\phi_{(l)}) + \frac{\hat{\gamma}}{r}\right) + \frac{\lambda}{r+\lambda+p}\left(V_L(\phi_{(0)}^{l-1}, R) + \frac{\hat{\gamma}}{r}\right) \\
&= V_L^{\Phi}(\phi_{(l-1)}, L) + \frac{\hat{\gamma}}{r}.
\end{aligned}$$

The claim thus follows for Case 2.

Now, let's return to Case 1. In this case, L's continuation value starting from $\phi_{(n-1)}$ is

$$\begin{aligned}
V_L(\phi_{(n-1)}, L) &= \frac{\pi_L(\phi_{(n-1)})}{\lambda+p+r} + \frac{p}{p+\lambda+r} V_L(\phi_{(n)}, L) + \frac{\lambda V_L(\phi_{(n-1)}, R)}{\lambda+p+r} \\
&= \frac{\pi_L(\phi_{(n-1)})}{\lambda+p+r} + \frac{p}{p+\lambda+r}\left(\frac{\pi_L(\phi_{(n)})}{\lambda+r} + \frac{\lambda V_L(\phi_{(n)}, R)}{\lambda+r}\right) + \frac{\lambda V_L(\phi_{(n-1)}, R)}{\lambda+p+r} \\
&> \frac{\pi_L(\phi_{(n)})}{\lambda+r} + \frac{\lambda}{\lambda+r} V_L(\phi_{(n)}, R) \\
&= V_L(\phi_{(n)}),
\end{aligned}$$

where the last inequality follows from the observations that $\pi_L(\phi_{(n)}) < \pi_L(\phi_{(n-1)})$ and $V_L(\phi_{(n)}, R) < V_L(\phi_{(n-1)}, R)$. This contradicts the fact that L's continuation value must be weakly increasing as he moves along his trajectory: $V_L(\phi_{(n)}, L) \geq V_L(\phi_{(n-1)}, L)$. Our claim thus holds in this case. ■

Lemma 7 *Define* $\underline{\zeta}_R = 1 + \frac{q}{p\left(1 - \left(\frac{p}{p+r}\right)^{\hat{\rho}_L + \hat{\rho}_R}\right)}$, *and define* $\phi_{(n)}$ *as follows:* $\phi_{(n)} \equiv (\underbrace{-1, -1, \cdots, -1}_{\hat{\rho}_L}, \underbrace{0, 0, \cdots, 0}_{n-\hat{\rho}_L})$ *for* $n > \hat{\rho}_L$, *and* $\phi_{(n)} \equiv (\underbrace{-1, -1, \cdots, -1}_{n})$ *for* $n \leq \hat{\rho}_L$. *If* $\zeta_R < \underline{\zeta}_R$, *then* R *will undo at any* $\phi_{(n)}$.

Proof. Borrowing from the proof of Lemma 4, we can calculate that (i) if $n \leq \hat{\rho}_L$,

R prefers to undo rather than extend at $\phi_{(n)}$ (equivalently, $V_0(\phi) < V_1(\phi)$) if $\zeta_R \leq 1 + \frac{2q}{(p-q)\left(1-\left(\frac{p}{p+r}\right)^{n+\hat{\rho}_R}\right)}$, and that (ii) if $n > \hat{\rho}_L$, R prefers to undo rather than extend at $\phi_{(n)}$ (equivalently, $V_0(\phi) < V_1(\phi)$) if $\zeta_R \leq 1 + \frac{q}{p\left(1-\left(\frac{p}{p+r}\right)^{\hat{\rho}_L+\hat{\rho}_R}\right)}$. Combining these conditions, we conclude that R will undo at any $\phi_{(n)}$ if $\zeta_R \leq 1 + \frac{q}{p\left(1-\left(\frac{p}{p+r}\right)^{\hat{\rho}_L+\hat{\rho}_R}\right)} = \underline{\zeta}_R$ ■

Lemma 8 *There exists $\underline{\zeta}_R > 1$ such that if $\zeta_R < \underline{\zeta}_R$, then for some $h \geq 0$, L's equilibrium policy trajectory $\Phi_h = (\phi_{(0)}, \phi_{(1)}, ..., \phi_{(h)})$ takes the following form:*

$$\phi_{(n)} \equiv (\underbrace{-1, -1, \cdots, -1}_{\hat{\rho}_L}, \underbrace{0, 0, \cdots, 0}_{n-\hat{\rho}_L})$$

for $n > \hat{\rho}_L$, and

$$\phi_{(n)} \equiv (\underbrace{-1, -1, \cdots, -1}_{n})$$

for $n \leq \hat{\rho}_L$.

Proof. Choose $\zeta_R < \underline{\zeta}_R$ from Lemma 8, so R will undo at any $\phi_{(n)}$. Before we go on, note that for all n, $\phi_{(n)}$ maximizes L's instantaneous payoff amongst all policies with complexity n. This fact (call it Fact A) will come in useful soon.

We will start by proving, by induction, the following claim (which we call Fact B): for any n,

$$V_L(\phi_{(n)}, R) \equiv \arg\max_{\phi}\{V_L(\phi, R) : \gamma(\phi) = n\}, \tag{6}$$

with the maximization strict for $n \geq 1$. First, note that this statement is trivially (albeit weakly) true for $n = 0$, because there is only one policy with zero complexity. Let us suppose that the induction hypothesis holds for n; then we will show that it holds for $n+1$ as well. We know that R will undo ϕ_{n+1} towards ϕ_n, so we have

$$V_{LR}(\phi_{(n+1)}) = \frac{\pi_L(\phi_{(n+1)})}{r+q} + \frac{qV_L(\phi_{(n)}, R)}{r+q}. \tag{7}$$

Consider an alternative policy ϕ' with the same complexity $n+1$. We seek to show that $V_L(\phi', R) < V_L(\phi_{(n+1)}, R)$. We have

$$V_L(\phi', R) = \frac{\pi_L(\phi')}{r+q} + \frac{qV_L(\phi'', R)}{r+q}. \tag{8}$$

where ϕ'' is the second policy on L's trajectory starting from ϕ'. Now, we know (via Fact A) that $\pi(\phi_{(n+1)}) \geq \pi(\phi')$. So, comparing equations (7) and (8), we simply need to show that $V_L(\phi_{(n)}, R) > V_L(\phi'', R)$ to verify our claim that $V_L(\phi_{(n+1)}, R) > V_L(\phi', R)$. There are two cases to consider. First, suppose R undoes at ϕ'. Then ϕ'' has complexity n; so $V_L(\phi_{(n)}, R) > V_L(\phi'', R)$ via the induction hypothesis. Second, suppose R extends at ϕ'. Let $\Phi'_R = (\phi'_{[0]}, \phi'_{[1]}, ..., \phi'_{[n'_R]})$ be R's equilibrium trajectory starting from $\phi' = \phi'_{[0]}$, and let $\Phi_R = (\phi_{[0]}, \phi_{[1]}, ..., \phi_{[n_R]})$ be R's equilibrium trajectory starting from $\phi_{(n+1)} = \phi_{[0]}$.

Notice that $(\phi_{[0]}, \phi_{[1]}, ..., \phi_{[n+1]}) = (\phi_{(n+1)}, \phi_{(n-1)}, ..., \phi_{(0)})$; i.e. starting from $\phi_{(n+1)}$, R retraces L's steps in reverse. Notice also that Φ_R contains weakly more steps than Φ'_R:

$$n'_R = \hat{\rho}_R + \hat{\rho}_L \leq n_R,$$

with strict inequality if $n + 1 > \hat{\rho}_L$. As a first step, we'll show that in a step-by-step comparison, Φ_R is more profitable than Φ'_R. Let $\rho_0 = \min\{0, n + 1 - \hat{\rho}_L\}$. For $k \leq n'_R$,

$$\pi^\rho_L(\phi'_{[k]}) = \left|\hat{\rho}_L - \rho(\phi'_{[k]})\right| = \left|\hat{\rho}_L - \rho(\phi'_{[0]}) + k\right| \geq \min\{k - \rho_0, 0\} \geq \pi^\rho_L(\phi_{[k]}),$$
$$\pi^\gamma_L(\phi'_{[k]}) = n + 1 + k \geq \pi^\gamma_L(\phi_{[k]}),$$

so $\pi_L(\phi'_{[k]}) < \pi_L(\phi_{[k]})$ (with strict inequality for $k \geq 1$). Further, for $n'_R < k < n_R$,

$$\pi^\rho_L(\phi'_{[n'_R]}) = \hat{\rho}_L + \hat{\rho}_R \geq \pi^\rho_L(\phi_{[k]}),$$
$$\pi^\gamma_L(\phi'_{[n'_R]}) = n + 1 + n_R > \max\{n, \hat{\rho}_R\} \geq \pi^\gamma_L(\phi_{[k]}),$$

so $\pi_L(\phi'_{[n'_R]}) < \pi_L(\phi_{[k]})$. With these results in hand, we'll show by induction that for $k \leq n'_R$, we have $V_L(\phi_{[k]}, R; \Phi_R) > V_L(\phi'_{[k]}, R; \Phi'_R)$. To initialize, note from our previous calculations that all policies on the trajectory Φ_R have $\pi_L(\phi_{[k]}) > \pi_L(\phi'_{[n'_R]})$, so $V_L(\phi_{[n'_R]}, R; \Phi_R) > V_L(\phi'_{[n'_R]}, R; \Phi'_R)$. Next, suppose the induction hypothesis holds for k. Then, by our induction hypothesis,

$$\begin{aligned} V_L(\phi_{[k-1]}, R; \Phi_R) &= \frac{\pi_L(\phi_{[k-1]})}{r + p} + \frac{pV_L(\phi_{[k]}, R; \Phi_R)}{r + p} \\ &> \frac{\pi_L(\phi'_{[k-1]})}{r + p} + \frac{pV_L(\phi'_{[k]}, R; \Phi'_R)}{r + p} = V_L(\phi'_{[k-1]}, R; \Phi'_R). \end{aligned}$$

It follows by induction that $V_L(\phi_{[0]}, R; \Phi_R) > V_L(\phi'_{[0]}, R; \Phi'_R)$, or equivalently that $V_L(\phi_{(n+1)}, R) > V_L(\phi', R)$, as we claimed.

Having established (6), we now consider a trajectory $\Phi'_m = (\phi'_{(0)}, \phi'_{(1)}, ..., \phi'_{(m)})$ with length m (note this means that $\phi'_{(m)}$ has complexity m) such that $\phi'_{(m)} \neq \phi_m$. Let $V_L^{\Phi'_m}(\phi'_{(k)})$ be the continuation value under this trajectory at policy $\phi'_{(k)}$. Also, let $V_L^{\Phi_m}(\phi_{(k)})$ be the continuation value under the trajectory Φ_m at policy $\phi_{(k)}$, with $k \leq m$. We claim that for all $k \leq m$, $V_L^{\Phi'_m}(\phi'_{(k)}, L) \leq V_L^{\Phi_m}(\phi_{(k)}, L)$ (with strict inequality for $k = 0$). This claim, once verified, establishes that starting from the origin, L prefers the trajectory Φ_m over any other Φ'_m, and thus establishes our lemma.

We proceed by induction. Start by comparing

$$V_L^{\Phi_m}(\phi_{(m)}, L) = \frac{\pi_L(\phi_{(m)})}{r + \lambda} + \frac{\lambda V_L(\phi_{(m)}, R)}{r + \lambda}, \text{ and}$$
$$V_L^{\Phi'_m}(\phi'_{(m)}, L) = \frac{\pi_L(\phi'_{(m)})}{r + \lambda} + \frac{\lambda V_L(\phi'_{(m)}, R)}{r + \lambda};$$

noting that $\pi(\phi_{(m)}) \geq \pi(\phi'_{(m)})$ (from Fact A) and that $V_L(\phi_{(m)}, R) > V_L(\phi'_{(m)}, R)$ (from

Fact B), we obtain $V_L^{\Phi_m}(\phi_{(m)}, L) > V_L^{\Phi'_m}(\phi'_{(m)}, L)$. Now, our induction step: suppose that, for given $k \leq m$, we have $V_L^{\Phi_m}(\phi_{(k)}) > V_L^{\Phi'_m}(\phi'_{(k)})$. Then we claim that this statement holds for $k-1$ as well. To show this, compare

$$V_L^{\Phi_m}(\phi_{(k-1)}, L) = \frac{\pi_L(\phi_{(k-1)})}{r+p+\lambda} + \frac{pV_L^{\Phi_m}(\phi_{(k)}, L)}{r+p+\lambda} + \frac{\lambda V_L(\phi_{(k-1)}, R)}{r+p+\lambda}, \text{ and}$$
$$V_L^{\Phi'_m}(\phi'_{(k-1)}, L) = \frac{\pi_L(\phi'_{(k-1)})}{r+p+\lambda} + \frac{pV_L^{\Phi'_m}(\phi'_{(k)}, L)}{r+p+\lambda} + \frac{\lambda V_L(\phi'_{(k-1)}, R)}{r+p+\lambda};$$

noting from Fact A that $\pi_L(\phi_{(k-1)}) \geq \pi_L(\phi'_{(k-1)})$, from Fact B that $V_L(\phi_{(k-1)}, R) > V_L(\phi'_{(k-1)}, R)$, and from our induction hypothesis that $V_L^{\Phi_m}(\phi_{(k)}, L) > V_L^{\Phi'_m}(\phi'_{(k)})$, it follows that $V_L^{\Phi_m}(\phi_{(k-1)}, L) > V_L^{\Phi'_m}(\phi'_{(k-1)})$. Thus our induction hypothesis holds, and so does our claim. ■

Proposition 9 *Suppose $\zeta_R > \bar{\zeta}_R$. Then L extends left at the origin if and only if $\zeta_L > \underline{\zeta}_L$. Otherwise, if $\zeta_L < \underline{\zeta}_L$, then L will stagnate at the origin.*

Proof of Proposition 9 Borrow $\bar{\zeta}_R$ from Lemma 4, so that for $\zeta_R > \bar{\zeta}_R$, player R always extends at any policy that is not R-ideal. Lemma 6 then states that L either stagnates or extends leftward at the origin. We show that there exists a $\bar{\zeta}_L$ such that for $\zeta_L > \bar{\zeta}_L$, L will extend at the origin.

Let V^R be the continuation payoff for player L at origin when R is in control. Then the value functions for player L from staying at origin ϕ_0, and at $\phi_{-1} = (-1)$ are, respectively,

$$V^L(\phi_0, L) \equiv \frac{\zeta_L \hat{\rho}_L + \lambda V^L(\phi_0, R)}{\lambda + r} \text{ and } V^L(\phi_{-1}, L) \equiv \frac{\zeta_L(\hat{\rho}_L + 1) - 1 + \lambda \frac{-\zeta_L(\hat{\rho}_L+1)-1+p\left(V^L(\phi_0,R)-\frac{2}{r}\right)}{p+r}}{\lambda + r}.$$

Note that
$$r\left(V^L(\phi_0, R) - \frac{2}{r}\right) < \underbrace{\zeta_L(\hat{\rho}_L + 1) - 1}_{\pi_L \text{ at } \phi=(-1)}.$$

Therefore, $V^L(\phi_{-1}, R) > \frac{-\zeta_L(\hat{\rho}_L+1)-1+\lambda\left(V^L(\phi_{-1},R)-\frac{2}{r}\right)}{\lambda+r}$, which implies
$(\lambda + r)\left(V^L(\phi_{-1}, L) - V^L(\phi_0, L)\right) > \zeta_L - \frac{2\lambda}{r} > 0$.

Lemma 9 *Suppose $\zeta_L > 2\lambda/r$. Then L extends at the origin, and does not stagnate if the policy is not R-ideal.*

Proof. Let $\phi_0^{'}$ be a policy that solely consists of left-biased rules such that $\rho\left(\phi'_{-1}\right) > \hat{\rho}_L$, where $\phi_{-1}^{'} = \phi_0^{'} \sqcup (-1)$. An analogous proof to Proposition 9 proves that

$$(\lambda + r)\left(V^L\left(\phi'_{-1}, L\right) - V^L(\phi_{,0}, L)\right) > \zeta_L - \frac{2\lambda}{r}.$$

Therefore, we have the required result. ■

Proof of Proposition ?? We consider the case where R is moderate; the case where L is moderate is covered in Proposition **??**.

Let Φ_L^* be the set of policies ϕ such that (i) $\rho(\phi) = \hat{\rho}_L^*$, and (ii) the last rule in ϕ is 0. Then, we show that there exists a $\underline{\zeta}_R^1$ such that if $\zeta_R < \underline{\zeta}_R^1$ then R prefers to undo exactly once than extend at any $\phi \in \Phi_L^*$. Also note that there exists a $\underline{\zeta}_R^2$ such that if $\zeta_R < \underline{\zeta}_R^2$ then R does any policies ϕ that only consists of -1 and $\gamma(\phi) \leq \hat{\rho}_L^*$. Then if $\zeta_R < \min\left\{\underline{\zeta}_R^1, \underline{\zeta}_R^2\right\}$, then L extends to L-ideal policy, and then either extends by adding 0, or stagnates. In either case, the kludge will not persist in the long run. ■

Proof of Proposition ??

First observe that since $\zeta_R > \bar{\zeta}_R$ L never extends by adding 1 or 0 on the equilibrium trajectory. Also, note that for any $m > \bar{m}(\zeta_L) \equiv (\zeta_L + 1)\hat{\rho}$, L does not extend to ϕ_m, where ϕ_m is the policy that consists of -1 rules with complexity of m. Fix an ζ_L^0, and define $\bar{m} = \bar{m}(\zeta_L)$. For any $m < \bar{m}$, there exists a $\underline{\zeta}_L^m$ such that L prefers to stagnates at ϕ_m than extending to ϕ_{m+1}. Therefore, if we define $\underline{\zeta}_L = \min_{m \in \{0,1,\cdots,\bar{m}\}} \underline{\zeta}_L^m$, then L stagnates at the origin if $\zeta_L < \underline{\zeta}_L$ and $\zeta_R > \bar{\zeta}_R \equiv \zeta_R\left(\tilde{\zeta}_L\right)$. ■

Proof of Proposition 6

Recall the proof of Proposition **??**. Then, L extends to ϕ_L^* and either extends by adding 0 or stagnates. The value functions at L's optimal policy ϕ_L^* and $\phi_L^* \sqcup (0)$ are

$$V_1 = \frac{\hat{\rho}_L + \lambda V^R}{r + \lambda} \text{ and } V_2 = \frac{\hat{\rho}_L - 1 + \lambda \frac{\hat{\rho}_L - 1 + pV^R}{r+p}}{r + \lambda}.$$

Note that $V^R < 0$ is decreasing in ζ_L and

$$(r + \lambda)(V_2 - V_1) = -1 + \lambda \frac{\hat{\rho}_L - 1 - rV^R}{r + p}.$$

Proof of Proposition 7 Fix r, and consider ζ_R and ζ_L that satisfy $\zeta_R > \bar{\zeta}_R(r) \equiv 1 + \frac{2q}{(p-q)\left(\frac{r}{p+r}\right)}$ and $\zeta_L > \tilde{\zeta}_L(r) \equiv \frac{2\lambda}{r}$. Then by Lemma 4 and Lemma 9, we know that L prefers extending to L's optimal policy ϕ_L^* and stagnates, than stagnating at the origin. So we only need to see when L wants to extends by adding -1 at ϕ_L^*.

The value functions for player L from staying at ϕ_L^*, and at $\phi = \phi_L^* \sqcup (-1)$ are, respectively,

$$V^L(\phi_L^*, L) \equiv \frac{\hat{\rho}_L + \lambda V^L(\phi_L^*, R)}{\lambda + r} \text{ and}$$

$$V^L(\phi_L^* \sqcup (-1), L) \equiv \frac{-(\zeta_L - \hat{\rho}_L + 1) + \lambda \frac{-(\zeta_L - \hat{\rho}_L + 1) + p\left(V^L(\phi_L^*, R) - \frac{2}{r}\right)}{p+r}}{\lambda + r}.$$

Below we show that there exists an $\underline{r}$ such that $r < \underline{r}$ implies $V^L(\phi_L^* \sqcup (-1), L) - V^L(\phi_L^*, L)$ is linear and increasing in ζ_L.

Now, observe that

$$\begin{aligned}&V^{L}\left(\phi_{L}^{*}\sqcup(-1),L\right)>V^{L}\left(\phi_{L}^{*},L\right)\\ \text{iff }&\lambda\left(\hat{\rho}_{L}-rV^{L}\left(\phi_{L}^{*},R\right)\right)>(p+r+\lambda)\left(\zeta_{L}+1\right)-\frac{2\lambda p}{r}.\end{aligned}$$

Moreover, note that

$$\begin{aligned}&\lambda\left(\hat{\rho}_{L}-rV^{L}\left(\phi_{L}^{*},R\right)\right)-(p+r+\lambda)\left(\zeta_{L}+1\right)-\frac{2\lambda p}{r}\\ &\geq\lambda\left(\hat{\rho}_{L}-rV^{L}\left(\phi_{L}^{*},R\right)\right)-(p+r+\lambda)\left(\zeta_{L}+1\right)-p\zeta_{L}\ (\text{by }\zeta_{L}>\frac{2\lambda}{r})\\ &\geq\lambda\left(\hat{\rho}_{L}-rV^{L}\left(\phi_{L}^{*},R\right)\right)-(2p+r+\lambda)\left(\zeta_{L}+1\right).\end{aligned}$$

Now define $\Gamma\left(\zeta_{L},r\right)\equiv\lambda\left(\hat{\rho}_{L}-rV^{L}\left(\phi_{L}^{*},R\right)\right)-(2p+r+\lambda)\left(\zeta_{L}+1\right)$. Next, observe that since $V^{L}\left(\phi_{L}^{*},R\right)$ is linear in ζ_{L}, so is $\Gamma\left(\zeta_{L},r\right)$. Therefore, $\frac{\partial\Gamma(\zeta_{L},r)}{\partial\zeta_{L}}$ is not a function of ζ_{L}, and

$$\begin{aligned}\lim_{r\to0}\frac{\partial\Gamma\left(\zeta_{L},r\right)}{\partial\zeta_{L}}&=\lim_{r\to0}\frac{\partial\lambda\left(\hat{\rho}_{L}-r\tilde{V}^{R}\right)}{\partial\zeta_{L}}-(2p+r+\lambda)\\ &=-2\lambda\hat{\rho}_{L}-(2p+r+\lambda),\end{aligned}$$

where the second equality follows from $\lim_{r\to0}\left(\hat{\rho}_{L}-rV^{L}\left(\phi_{L}^{*},R\right)\right)=-2\hat{\rho}_{L}\left(\zeta_{L}+1\right)$. Since $\lambda>2p$ and $\hat{\rho}_{L}\leq-1$, $\lim_{r\to0}\frac{\partial\Gamma(\zeta_{L},r)}{\partial\zeta_{L}}>0$. Therefore, there exists an $\underline{r}$ such that for all $r<\underline{r}$, $\zeta_{L}'>\zeta_{L}$ and $\Gamma\left(\zeta_{L},r\right)>0$, imply $\Gamma\left(\zeta_{L}',r\right)>0$.

Now that we have established that $V^{L}\left(\phi_{L}^{*}\sqcup(-1),L\right)-V^{L}\left(\phi_{L}^{*},L\right)$ is linear and increasing in ζ_{L} for any $r<\underline{r}$. Therefore, for any $r<\underline{r}$, there exists a $\hat{\zeta}_{L}(r)$ such that $\zeta_{L}>\hat{\zeta}_{L}(r)$ if and only if $V^{L}\left(\phi_{L}^{*}\sqcup(-1),L\right)>V^{L}\left(\phi_{L}^{*},L\right)$. Therefore, if we define $\bar{\zeta}_{L}(r)=\max\left\{\bar{\zeta}_{R}(r),\hat{\zeta}_{L}(r)\right\}$, we have the required result.■

Proof of Proposition 8

First, we show that there exists an $\bar{h}$ such that for any p and q, L does not extend any policy that is more complex than $\bar{h}$, irrespective of R's strategy. To see this, consider a policy ϕ such that

$$\phi=\left(\underbrace{-1,-1,\cdots,-1}_{\hat{\rho}_{L}},\underbrace{0,0,\cdots,0}_{h-\hat{\rho}_{L}}\right).$$

That is, ϕ is L-ideal and has complexity h. Therefore, L's instantaneous payoff from any policy that is at least as complex as ϕ is strictly lower than that of at ϕ, i.e., $\pi_{L}(\phi)=-\left(h-\hat{\rho}_{L}\right)$. Now let $\tilde{\pi}_{L}$ be L's instantaneous payoff at R's optimal policy. If L extends to ϕ in an equilibrium, then L's strategy has to satisfy $-\left(h-\hat{\rho}_{L}\right)>\tilde{\pi}_{L}=\left(2\zeta_{L}+1\right)\hat{\rho}_{L}$. This proves that L does not extend to any policy with complexity of more than $\bar{h}\equiv2\zeta_{L}\hat{\rho}_{L}$.

Next, consider a policy ϕ. Let ϕ^{E} be the policy that an R-favoured rule is added to ϕ. Similarly, let ϕ^{U} be the policy without the last rule in ϕ. Then the value functions of extending and undoing at ϕ are, respectively,

$$\frac{\pi_R(\phi) + pV\left(\phi^E\right)}{p+r} \text{ and } \frac{\pi_R(\phi) + qV\left(\phi^U\right)}{q+r}$$

Note that $\lim_{\chi \to 0}\left(V\left(\phi^U\right) - V\left(\phi^E\right)\right) \geq 1/r$. Therefore, as $\chi \to 0$, $\frac{\pi_R(\phi)+qV\left(\phi^U\right)}{q+r} - \frac{\pi_R(\phi)+pV\left(\phi^E\right)}{p+r} \to \alpha$ for some $\alpha \geq 1/r$. Therefore, for any $\chi < \underline{\chi}(\phi)$, R undoes at ϕ. Thus if we define $\underline{\chi} \equiv \min_{\phi \in \Phi_h} \underline{\chi}(\phi)$, where $\Phi_{\bar{h}}$ is the set of policies that has length $k \leq h$, then R undoes any policy that L reaches with a positive probability on the equilibrium. This proves the first part of the proposition.

Similarly, note that

$$\begin{aligned}
&\lim_{\chi \to \infty}\left(\frac{\pi_R(\phi) + pV\left(\phi^E\right)}{p+r} - \frac{\pi_R(\phi) + qV\left(\phi^U\right)}{q+r}\right) \\
&\geq \lim_{\chi \to \infty}\left(\frac{\pi_R(\phi) + p\left(\frac{\pi_R(\phi)+(\zeta_R-1)}{r}\right)}{p+r} - \frac{\pi_R(\phi) + q\left(\frac{\pi_R(\phi)+(\zeta_R+1)}{r}\right)}{q+r}\right) \\
&= \lim_{\chi \to \infty}\left(\frac{r(p-q)\zeta_R - (2qp + (p+q)r)}{r(p+r)(q+r)}\right) \\
&= \frac{\lim_{\chi \to \infty}((p-q)\zeta_R - (p+q))}{r^2} > 0.
\end{aligned}$$

Therefore, there exists a $\overline{\chi}(\phi)$ such that $\chi > \overline{\chi}(\phi)$ implies R extends at ϕ. If we define $\overline{\chi}(\phi) \equiv \max_{\phi \in \Phi_{\bar{h}}} \overline{\chi}(\phi)$, then we have the required result. ■

References

Alesina, A., and A. Drazen (1991): "Why Are Stabilizations Delayed?," *American Economic Review*, 81(5), 1170–88.

Callander, S. (2011): "Searching for Good Policies," *American Political Science Review*, 105(4), 643–662.

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

Ellison, G., and R. Holden (2013): "A Theory of Rule Development," *Journal of Law Economics and Organization*.

Ely, J. C. (2011): "Kludged," *American Economic Journal: Microeconomics*, 3(3), pp. 210–231.

Lippman, S. A. (1975): "On Dynamic Programming with Unbounded Rewards," *Management Science*, 21(11), pp. 1225–1233.

——— (1976): "Countable-State, Continuous-Time Dynamic Programming with Structure," *Operations Research*, 24(3), pp. 477–490.

Morris, S., and S. Coate (1999): "Policy Persistence," *American Economic Review*, 89(5), 1327–1336.

Teles, S. (2013): "Kludgeocracy in America," *National Affairs*, 17, 97–114.