From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Economics

2021

Essays in Dynamic Cooperation: Bargaining and Contribution Games

Parihar, Parth

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp016682x7099

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

ESSAYS IN DYNAMIC COOPERATION: BARGAINING AND CONTRIBUTION GAMES

Parth Singh Parihar

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY

THE DEPARTMENT OF

ECONOMICS

ADVISER: LEEAT YARIV

SEPTEMBER 2021

© Copyright by Parth Singh Parihar, 2021.

All rights reserved.

Abstract

In many economic and political settings, agents with non-aligned interests must work together to enact mutually beneficial change. My dissertation consists of three chapters that analyze problems that combine features of bargaining and contribution games. In particular, I examine how institutional rules, the size of teams, and the attitudes of agents affect dynamic cooperation.

In Chapter 1, I analyze a repeated bargaining environment where the previously negotiated agreement becomes the present status quo. Agreements' efficiency decays in the absence of revision, giving both parties an incentive to re-negotiate. I study how limited foresight— or the inability of agents to account for the full dynamic linkage from one agreement to the next— affects behavior. I show that limited foresight is a key driver of gridlock. Moreover, I demonstrate that while limited foresight and inter-temporal discounting both overweight imminent outcomes, these have the *opposite* effects on equilibrium. I solve the model in the case of legislative bargaining and relate its findings to observed phenomena in policy-making.

In Chapter 2, Matias Iaryczower, Santiago Oliveros, and I study sequential contributions to public goods. We consider a model where agents with different valuations for the good are selected at random to contribute in each period. Agents face classical *free riding* incentives in preferring to offload the cost of contributing to other players. We show that if the project is sufficiently large, the unique equilibrium of the model displays *endogenous contribution cycles*, in which agents of different types alternate making gradual contributions towards completion of the project.

Finally, in Chapter 3, I explore a setting that combines features of both bargaining and contribution games. Two players cooperate to complete a joint project, wherein each of several periods, one agent— the project leader— is responsible for advancing the project at personal cost. Project leaders are choose not only the extent of progress they enact but also the *direction* in which the project moves, over which their preferences diverge with their counterparts. I show that increasing the incumbency of an efficient project leader *decreases* her productivity while the opposite is true for an inefficient contributor.

Acknowledgements

My graduate work in Princeton, culminating in this dissertation, has been both a tremendous joy as well as a challenge that continually pushed me to dig deeper within myself. I would therefore be remiss if I did not thank those many people who made it such a joy and also gave me the strength to surmount these challenges.

I am deeply indebted to my extraordinary advisor, Leeat Yariv, whose dedication is perhaps unparalleled. I feel unimaginably fortunate to have had you as a role model for what an academic should be. Thank you for making me feel as though I belong in this profession. I also must thank you for your patient and forgiving attitude— yes, in encouraging me to push forward in my work, but also in overlooking the unfortunate incident of my dog peeing on your brand new sofa. Jokes aside, your opinion will always mean the world to me.

To Wolfgang Pesendorfer, thank you for being there throughout all the small, but important, milestones of my academic career here at Princeton. You taught my first (ECO 310) and last (ECO 520) economics courses— seven years apart— and helped me learn the ropes of research by advising my third year paper. To Matias Iaryczower, thank you for your enthusiasm, patience, and sense of humor, which has made working with you a wonderful treasure that I hope to continue and can only dream of repaying in kind.

I am incredibly grateful to participants in the Micro-economic Theory and Political Economy lunch workshops and seminars for sitting through each of my talks, both good and bad. Thank you to Faruk Gul, Sylvain Chassang, Can Urgun, Alessandro Lizzeri, Gleason Judd, German Gieczewski, and Thomas Romer for the many conversations and suggestions that have greatly improved my work. To Alessandro in particular, the conversation we had over the phone after my final lunch seminar presentation was immensely helpful and reassuring.

To Laura Hedden, I'm sorry that our interactions were always on occasions that I needed help on an administrative topic. Thank you for being so incredibly kind and empathetic and making every step of Princeton so much easier (including this last one). To Ellen Graf, thank you for all your administrative support during the stressful job market months.

My graduate studies at Princeton would not have been a joy were it not for the many, many friends that I've made along the way. Thank you to Emily Cuddy, Pëllumb Reshidi, George Langhans, Stefano Baratuche, and Bruno Baranek for board game nights that I always looked forward to. To my office mates— Franz Ostrizek, Mu Zhang, Simon Lazarus, and João Thereze— thank you for all your helpful advice and spontaneous interest in my work. I will miss our big office a lot. Thank you to Dr. Indu Viswanathan for always being someone I could come to with anything, for her encouragement, and for being such a courageous example. To the 2Dickinson Street Co-Op (2D), thank you for giving me a lifelong love of cooking and for being such a loving community for the five years I was fortunate enough to be a part of it. Finally, a huge shout to to my Hindu Students Council family— Nikunj, Arnav, Sohini, Ravi, Priya, Riddhi, and many others— thank you.

This dissertation is a culmination of not only six years of graduate studies but the many that preceded it. Foremost, Dilip Abreu, thank you for being the one who convinced me to study economics as a junior at Princeton. Thank you also to my many K-12 teachers who have shaped my intellectual curiosity in ways I don't yet even fully understand. To Mrs. Boyajian, Mr. Monti, Mrs. Strauch, Mrs. DeAngelis, Ms. Pancari, Mr. Barnshaw, Mrs. Carchidi, Mr. Morris, Mr. Wright, and Mr. Dempsey— thank you.

Most of all, I owe all the gratitude in the world to my family— Mama, Papa, Prachi, Siddharth, Baba and Dadi ji, and of course, Chetak. You've loved and supported me in more ways that I can recount throughout my whole life. Mama and Papa, you've always put Prachi and my interests above your own and did whatever you could to give us a lifelong thirst for learning. Prachi, you were the big sister role model that any little brother would want to have. This work is dedicated to you three and all the educators who made it possible.

To Mama, Papa, Prachi, and

all of my teachers that have made this possible.

Contents

Abstract iii

Acknowledgements iv

1 Limited Foresight and Gridlock in Bargaining **1**

1.1 Introduction 1

1.1.1 Related Literature 6

1.1.2 Structure of the Paper 8

1.2 Model 8

1.2.1 Basic Set-Up 8

1.2.2 Agreements and Foresight Horizon 11

1.2.3 Strategies and Equilibrium Concept 13

1.3 Determinants of Gridlock 14

1.3.1 The Incidence of Gridlock 16

1.3.2 Infinite Foresight Horizon 18

1.4 Legislative Bargaining 19

1.4.1 The Setting 20

1.4.2 Equilibrium Characterization 22

1.4.3 Policy Dynamics and Incrementalism 27

1.4.4 Uncertainty About the Median Voter 32

1.5 Conclusion 35

1.5.1 Extensions 35

1.5.2 Concluding Remarks . . . 38
References . . . 40
1.A Omitted Proofs for Section 3 . . . 43
1.B Omitted Proofs for Section 1.4 . . . 48
1.B.1 Proofs for Section 1.4.2 . . . 49
1.B.2 Proofs for Section 1.4.3 . . . 56
1.B.3 Proofs for Section 1.4.4 . . . 57
1.C A Micro-Foundation for Limited Foresight . . . 64
1.C.1 Endogenizing Foresight in Equilibrium . . . 64
1.C.2 Discussion . . . 65

2 Endogenous Contribution Cycles **68**
2.1 Introduction . . . 68
2.2 Related Literature . . . 72
2.3 Model . . . 75
2.4 Results . . . 76
2.4.1 Small Projects . . . 77
2.4.2 Large Projects . . . 86
2.5 Conclusion . . . 99
References . . . 102
2.A Proofs for Chapter 2 . . . 103
2.A.1 Proof of Lemma 6 . . . 103
2.A.2 Proof of Proposition 8 . . . 104
2.A.3 Proof of Proposition 9 . . . 115
2.A.4 Proof of Proposition 10 . . . 116
2.A.5 Proof of Theorem 3 . . . 117
2.A.6 Proof of Theorem 11 . . . 120

3 When Incumbency Limits Productivity **122**

3.1 Introduction . . . 122

3.1.1 Literature Review . . . 124

3.1.2 Paper Outline . . . 125

3.2 Model . . . 126

3.2.1 Timeline . . . 127

3.2.2 Strategies and Equilibrium . . . 129

3.3 Preliminaries and the Finite Deadline Game . . . 131

3.3.1 Analytic Framework . . . 131

3.3.2 Equilibrium Characterization . . . 133

3.4 Main Results . . . 139

3.4.1 Incumbency and Productivity . . . 139

3.4.2 Application: Government Turnover and Spending . . . 141

3.5 Conclusion . . . 143

References . . . 144

3.A Omitted Proofs for Section 3 . . . 146

3.B Omitted Proofs for Section 4 . . . 150

Chapter 1

Limited Foresight and Gridlock in Bargaining*

1.1 Introduction

In bargaining environments, agents often inherit status quo agreements mutually regarded as sub-optimal. For instance, geopolitical or technological developments may render existing treaties between rival countries insufficient to tackle a common challenge, such as climate change. Within cartels consisting of producers with heterogeneous capacities, exogenous demand shocks can make existing supply arrangements dominated for all members. Likewise, in legislative settings, different parties often concur on the need for reform on a given issue— for instance, immigration within the modern-day U.S. polity— despite having drastically different visions on how to enact it.

Yet, despite this powerful common incentive for renegotiation, parties often encounter inefficient delay, or *gridlock*, in arriving at new agreements. This was the case, for example, at the 2009 Copenhagen Climate Change Conference, in which nations failed to generate an international climate accord to replace the floundering Kyoto Protocol; this task was left until 2015, in Paris. In March

*I am immensely grateful to Leeat Yariv, Pietro Ortoleva, Alessandro Lizzeri, Wolfgang Pesendorfer, Matias Iaryczower, Antonio Niccolo, Sylvain Chassang, Adam Meirowitz, Gleason Judd, and Steven Callander for their feedback and encouragement. I would also like to thank audiences from the University of Manchester and the Princeton political economy and microeconomic theory seminars for their helpful comments.

2020, as energy demand plummeted worldwide in response to the COVID-19 pandemic, negotiations over further supply cuts amongst OPEC (Organization of Petroleum Exporting Countries) members broke down. Here, delay was costly in the most direct sense, creating a glut of crude oil in the interim and hurtling prices to record lows.[1] Moreover, in the recent past, legislative gridlock has become an alarming norm in the United States. For example, a recent study found that Congressional gridlock on legislation dealing with the country's most "salient issues" had increased from 27% in 1947-48 to over 70% in the 2011-12 Congress.[2]

Gridlock is a mystifying feature of negotiations, since the canonical models of bargaining suggest that agreement should be immediate— e.g., Rubinstein (1982) and Baron and Ferejohn (1989). After all, if surplus is available from arriving at an agreement, and bargaining parties have policy instruments available to divide that surplus arbitrarily, why *don't* they?

This paper is concerned with two central questions. The first is— *under what circumstances does gridlock arise?* I demonstrate that gridlock can occur as a strategic consequence of two distinct features of any bargaining environment. First, as in the situations described above, agreements continue to remain in effect and grow progressively obsolete over time, until a successful renegotiation is reached. I call this *policy decay*. Second, negotiators consider only a finite number of future revisions of policy from the entire chain of policy consequences that follow from enacting a new agreement in the current period. I refer to this as *limited foresight*.

I then direct attention to the specific setting that signposts this paper, *legislative bargaining*. Here, we ask— *what are the consequences of limited foresight and gridlock on policy dynamics and electoral politics?* Comparing features of equilibrium across different levels of agent foresight and other parameters of the model, I relate equilibrium behavior to phenomena in public policy-making.

Our focus on studying the effects of limited foresight on bargaining outcomes is motivated by a central, under-explored insight of many observers. Mansbridge and Martin (2013) note that,

[1] Meredith, Sam. "OPEC+ Fails to Agree on Massive Supply Cut, Sending Crude Prices to 2017 Lows," *Consumer News and Business Channel (CNBC).* 6 March 2020. https://www.cnbc.com/2020/03/06/opec-meeting-coronavirus-weighs-on-oil-demand-as-oil-prices-fall.html.

[2] Binder, Sarah. "Polarized We Govern?" *Brookings Institute.* 27 May 2014. https://www.brookings.edu/research/polarized-we-govern/. Here, salience is defined by coverage in the mainstream press at a high rate.

> as negotiators, we may suffer from a *time myopia* that diminishes our capacities for long-term thinking, our material incentives for "saving for a rainy day," and our abilities to consider second and third-order effects. ... Costs and benefits often are structured to exacerbate the problem because few want to make short-term investments in exchange for risky, long-term rewards.[3]

That is, due to rationality constraints, decision makers often exhibit a form of present bias that leads them to underweight future policy consequences the more downstream these are from agreements hatched in the present. Agents focus instead on short-term policy gains, especially when there is uncertainty over long-run outcomes. Our notion of limited foresight captures precisely this long-run utility framework. This behavior has especially been noted in the legislative context,[4] further motivating this avenue of exploration.

We begin by studying the incentives underlying gridlock, in general, through a dynamic model of two-party bargaining in which status-quo policies are endogenous. In each period, two agents— L and R— inherit a status-quo policy characterized by its efficiency and underlying partisan lean. These characteristics are not co-dependent; a policy can be efficient and R-favoring, or inefficient and egalitarian, for example. One of the agents is selected to serve as the *incumbent* in the current period. The incumbent makes a proposal, and the other agent— the opposition— either accepts or rejects it. If accepted, the proposal is enacted as policy; otherwise, the status-quo remains in effect. The utility associated with the outcome accrues to all. As the game moves to the next period, utilities are discounted by an exponential discount factor $\delta \in (0, 1)$ and the policy enacted last period is inherited, but with a depreciated efficiency. This policy decay captures the way in which technological innovation, demographic shifts, and geopolitical re-alignments, among other changes, can leave existing agreements less able to serve their own intended purpose.

[3] Mansbridge and Martin, *Report of the American Political Science Association's Task Force on Negotiating Agreement in Politics* (2013), p. 122.

[4] See Lindblom (1959) and Hayes (2017). Lee (2016), p. 73 provides a colorful metaphor— "when the battle for party control is more competitive, a minority party's *time horizon shortens*, and its long game and short game collide. Members of such a minority party are more likely to throw bombs, sharpen the differences, and let substantive policy making wait for the day when they're no longer just making contributions at the margins."

Importantly, as the efficiency of the status-quo policy decreases, policy revision grows more attractive and urgent for all players. For instance, during crises, there is widespread demand for action to ameliorate the situation. The incumbent thus gains leverage over the opposition in acquiring favorable concessions in exchange for immediate action. At the same time, in the absence of agreement, since the next period's incumbent will have *even more* leverage as a result of further policy decay, agreement incurs an opportunity cost associated with foregoing a possibly even more favorable agreement in the next period. Agreement is reached only when the gains in efficiency from action outweigh this opportunity cost for both parties. Thus, any equilibrium of the game is characterized by *cut-off strategies* (Proposition 1), where agreement is reached only when the efficiency of status-quo policy falls below a unique threshold. With limited foresight, this threshold is *below* the fully-efficient benchmark, resulting in gridlock (Corollary 2).

When bargaining agents exhibit more limited foresight, the opportunity cost grows larger relative to potential gains in efficiency, exacerbating gridlock. As an extreme example, if agents behave as though there is only "one bite at the apple" of reform, the opportunity cost grows more pronounced. The impending revision represents *the only relevant* opportunity to instantiate a favorable agreement. On the other hand, since only the current revision holds any relevance, potential efficiency gains from further downstream agreements play no role in decision-making.

In contrast, when the foresight horizon T of bargaining agents is larger, they incorporate more future agreements into their objective functions, acknowledging future opportunities to mold policy to favor their own interests. Thus, each agent signals a willingness to forfeit immediate gains for future ones, attaching relatively less import to the partisan lean of current policy. A greater foresight horizon therefore reduces gridlock (Proposition 2 (a)). In fact, gridlock is a persistent feature of equilibrium if and only if foresight is limited (Proposition 3).

We consider also the effect of temporal impatience on gridlock. Note that temporal impatience and limited foresight both overweight imminent outcomes. Yet, one of the central contributions of this paper is to underline that these have the *opposite* effect on the prevalence of gridlock— in particular, a larger inter-temporal discount factor δ increases gridlock (Proposition 2 (b)). This is

because δ measures incentives to put off ideological concessions for *future* periods; hence, the cost (in lost surplus) of waiting until a future period to enact reform, is *smaller* the *larger* is δ.

Next, we pivot to studying the consequences of limited foresight and gridlock in legislative bargaining. This setting inherits all the features of the general model, where L and R now take on the meaning of left- and right-wing political parties. We introduce an electorate modeled as a continuum of voters, each with her own preference over the partisan lean of policy. Political parties do not have the ability to *ex ante* commit to specific proposals. Rather, each voter infers what legislative outcome *would be enacted* were either party elected, and votes for the outcome she prefers given her ideological preference (type). The incumbent is the party that obtains the majority vote share in each period. The identity of the median voter each period is stochastic and *ex ante* unknown to agents so that proposal power is both fully endogenous and non-deterministic.

I show there exists and solve for the unique equilibrium of the model when the median voter is distributed uniformly over the ideological continuum (Proposition 4). I demonstrate that the expected partisan lean of the next agreement is positively correlated with that of the current status quo (Corollary 3). Intuitively, the party that obtains greater utility from the outside option (status quo) has greater leverage over the next enacted agreement. This correlation also establishes that each party will propose policy on the extrema of the set of mutually acceptable agreements in order to maximize long-run utility. Since players are more willing to forego immediate policy gains as T increases, the distance between parties' proposals is increasing in the foresight horizon T (Proposition 5). Thus, greater foresight implies greater policy *divergence*.

Moreover, I study the policy dynamics of long-run equilibrium in the model. For any horizon T, the partisan lean of policy converges on the path of equilibrium play to a unique distribution symmetric about the expected median voter (Proposition 6). Using simulations, I show that the shift in the partisan lean of policy from one renegotiation to the next is smaller the less foresighted politicians are. This finding provides an endogenization within a formal bargaining model of Lindblöm's influential theory of incrementalism— the idea that politicians tend to enact policies similar to the status quo. Finally, I extend the results from the case of an i.i.d. uniformly distributed

median voter to show these hold for a broader family of distributions where the degree of uncertainty (variance) over the median voter's position varies (Proposition 7). I additionally show that as this variance decreases, the midpoint of the proposals shifts towards the expected median voter type, $\frac{1}{2}$. However, the proposals of both parties typically move in the *same* direction, rather than converging. This distinguishes behavior under our model from standard results in electoral spatial competition—e.g., Wittman (1983).

In the remainder of this Section, I discuss related literature in context, and delineate the structure of the paper.

1.1.1 Related Literature

This paper connects the literature on dynamic bargaining with an endogenous status quo with that on dynamic electoral accountability. The connection arises from viewing policy-making as a bargaining process that (1) considers how current decision-making influences future outcomes, and (2) allows voters to allocate proposal power, as is explored in Section 1.4.

The paper is therefore most closely related to papers on dynamic legislative bargaining with an endogenous status-quo. Dziuda and Loeper (2016) consider a model with two parties and policy instruments in which a utility-relevant, continuous state changes each period. Parties' ideological preferences are not aligned. As a result, an agent who would favor changing the policy, given the state alone, worries that the policy will remain in effect in future states *less favorable* to her (but not to her counterpart). Hence, the threshold for agreement is increased beyond that which would be Pareto efficient. The key insight of gridlock emerging as a result of future policy's *path dependence* on current policy is also present in Fearon (1996), Austen-Smith et al. (2019) and Acharya and Ortner (2020). I build on this work by considering a model in which proposal power

has a non-trivial impact on policy dynamics and is determined endogenously.[5] My framework is novel in that consideration of future policy consequences can help *engender* agreement.

Callander and Martin (2017) present a model of repeated bargaining between a fixed incumbent and opposition, in which there is a similar notion of policy decay. Callander and Martin (2020) introduce gridlock into this set-up by imposing that *if an incumbent is unable to enact policy, she loses power* in the next election. Viewed through the efficiency gains-opportunity cost framework introduced before, this exogenous, retrospective assumption makes the opportunity cost of acquiescing to a proposal very costly for an opposition party, since it is *sure* to be in power the next period if it rejects the proposal, when it can negotiate more favorably. A less stark version of this electoral dynamic drives gridlock in Ortner (2017). In my model, with risk-netural agents, I find that this effect vanishes without limited foresight. Other related papers in this literature include Kalandrakis (2004), Duggan and Kalandrakis (2012), and Anesi and Duggan (2018).

More broadly, this paper sits within the literature on dynamic bargaining with frictions. For instance, Ali (2006), Yildiz (2011), Ortner (2013), among others, show that optimism— irrationally high expectations about future proposal power— can lead to gridlock. Contextually, optimism increases the opportunity cost of *both* agents from accepting a proposal, while potential gains in efficiency remain constant. However, when agents have common knowledge about optimism, Yildiz (2003) shows that agreement is immediate in a bilateral setting. I show when bargaining is bilateral and each agents exhibits limited foresight, common knowledge of that limited foresight is *not* an impediment to gridlock. While models of optimism in bargaining typically consider only a single agreement, I focus here on the dynamic linkage of policy across time and its effects on policy dynamics and gridlock.

Finally, by endogenizing proposal power via elections,the current paper connects bargaining to the literature on dynamic electoral accountability— e.g., Duggan (2000), Bai and Lagunoff (2011),

[5]Note that if there are only two policy instruments, as in Dziuda and Loeper (2016), proposal power plays no role when both parties must agree to enact legislation. Extensions with more than one instrument continue to underrate the proposal power of the incumbent by implementing a step-by-step "consensus procedure" (see, Secion IV of Dziuda and Loeper (2016)). Of the papers that explore path-dependence as a determinant of gridlock, only Fearon (1996) includes electoral considerations.

and Banks and Duggan (2008). These papers analyze scenarios in which incumbents face an explicit trade-off between instituting preferred policies and maintaining power in the future. Each finds that this trade-off helps moderate incumbents' policy proposals. In the model I analyze, a similar effect emerges from *both* the electoral incentives as well as the constraints imposed by the bargaining environment.

1.1.2 Structure of the Paper

Section 1.2 presents the general model. Section 1.3 discusses all of the main results on gridlock in the general setting. Section 1.4 presents a more specific model that is attuned to the setting of legislative bargaining with endogenous elections. Results of simulations are also provided to help visualize policy dynamics and provide key insights. I also examine the role that uncertainty over the nature of the electorate plays in determining equilibrium policies. Section 1.5 provides a brief discussion of extensions of the model and finally concludes.

1.2 Model

1.2.1 Basic Set-Up

We consider a dynamic bargaining game between two parties (agents), L and R. Feasible outcomes $p \in P := [0,1]^2$ are described both by their efficiency $e \in [0,1]$ as well as the ideology of the underlying agreement, $a \in [0,1]$. e captures how well the the corresponding outcome objectively responds to the underlying, mutual problem it is meant to solve in the present, while a describes its partisan lean (how much it favors R). Each agent has an ideal point x_i corresponding to the ideological dimension of outcomes, such that bargaining agent i with ideal point x_i accrues utility

$$u_{x_i}(p) = e \cdot (1 - |x_i - a|)$$

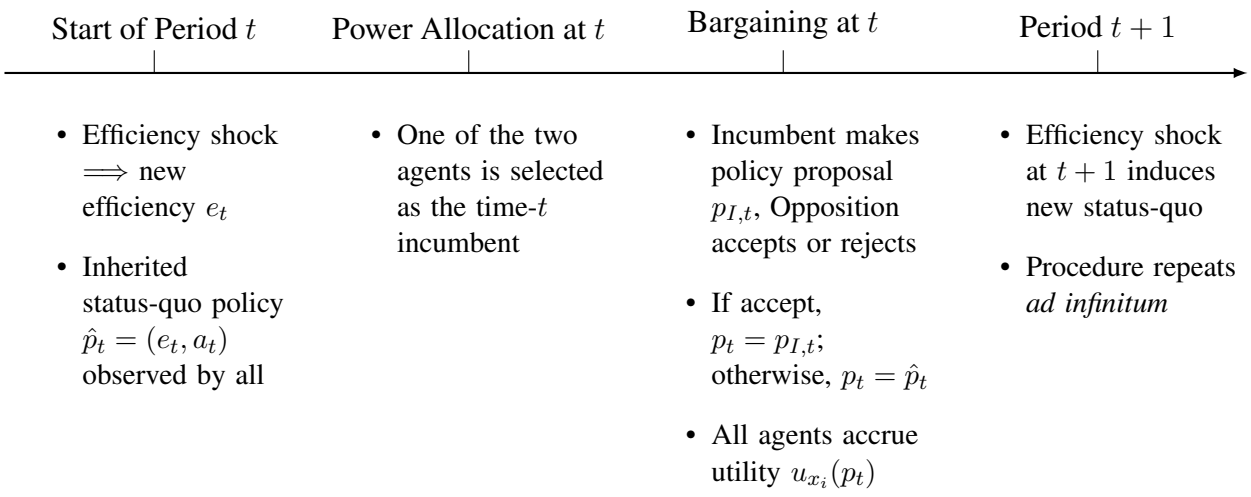


Figure 1.1: The Timeline of Within-Period Events

in each period where outcome $p = (e, a)$ materializes. I normalize the ideal points of the bargaining agents to $x_L = 0$ and $x_R = 1$, respectively. While the utility of agents is increasing in the efficiency of any outcome, they disagree on its partisan component. In fact, the normalization allows us to view each outcome (e, a) as a surplus of size $e \in [0, 1]$ which is divided so that R receives a fraction a of the surplus and L the residual portion $1 - a$.

I assume an infinite-horizon, repeated bargaining game as follows. At the start of period $t \in \mathbb{N}$, there is a publicly observable, stochastic shock $\tilde{\beta}_t$ to the efficiency of policy. If the outcome implemented in period $t-1$ is $p_{t-1} = (e_{t-1}, a)$, then the status-quo outcome inherited by all players in period t is given by

$$\hat{p}_t = (\tilde{\beta}_t \cdot e_{t-1}, a).$$

We make the following assumption on the sequence of shocks $\{\tilde{\beta}_t\}$—

Assumption 1. *The stochastic shocks $\tilde{\beta}_t$ are identically and independently distributed for all $t \in \mathbb{N}$ according to a nonatomic distribution F on $[0, 1]$. We denote $\overline{\beta}$ the expected value of $\tilde{\beta}_t$.*

Hence, the outcome that materializes at time t in the absence of policy revision has the same partisan lean as the previous period's, but is less efficient. We interpret the shock as the effect of idiosyncratic developments that unfold while a policy stays in effect that erode its efficiency.

Once $\hat{p}_t$ materializes and is observed by all agents, one of the agents is selected[6] as the time-t *incumbent* ($I \in \{R, L\}$), with the other player the *opposition*.[7] Both the incumbent and the opposition then bargain over the period-t outcome, p_t, which is determined as follows: the incumbent proposes an outcome $p_{I,t} = (e_{I,t}, a_{I,t})$. After observing the proposal, the opposition chooses whether to oppose ($o_t = 1$) or accept ($o_t = 0$) it. The outcome at time t is given by

$$p_t = (1 - o_t)p_{I,t} + o_t \cdot \hat{p}_t.$$

This set-up captures the reality that many institutions of collective choice give stakeholders power to veto a proposal made by one or more of its constituent members.[8] The timeline of the model discussed so far is summarized in Figure 1.1 above. The structure of the model as presented allows us to define key terms that will prove useful in introducing the specific details that remain.

Definition 1.

1. *An **agreement** a **is in place** in periods $\underline{t}, \ldots, \overline{t}$ if $p_t = (e_t, a)$ for some e_t for every $t \in \underline{t}, \ldots, \overline{t}$, $o_t = 1$ for every $t \in \underline{t} + 1, \ldots, \overline{t}$, and $o_t = 0$ for $t = \underline{t}$ and $t = \overline{t} + 1$*

2. *Suppose an agreement $a \in [0, 1]$ is in place in periods $t, \ldots, t + N$. The **efficiency life** of agreement a is*

$$\eta^a = \sum_{i=0}^{N} \delta^i e_{t+i},$$

where $\delta \in (0, 1)$ is the common, exponential rate at which players discount future outcomes.

Hence, an *agreement* refers to the partisan lean of an outcome and remains in place until a revision of the policy is arrived at through the repeated bargaining procedure described above.[9] The

[6] As we impose Markov strategies (see Section 3), we will assume that the mechanism that selects the incumbent depends only on the status quo and the foresight horizon T of players. Formally, a *mechanism* π is a mapping $\pi : P \times \mathbb{N} \to \Delta(\{L, R\})$.

[7] These names are interchangeable with proposer and responder and are denoted to conform to the motivating example of legislative bargaining.

[8] In particular, this captures the reality of democratic institutions. Examples of such institutions include a higher threshold of support for a constitutional amendment, the filibuster in the U.S., a bicameral legislature, etc.

[9] Note that the requirement that $o_t = 1$ for every period subsequent to the one in which the agreement is initialized means that a revision of the outcome ($o_t = 0$) which induces the same partisan lean is considered to be a *different* agreement.

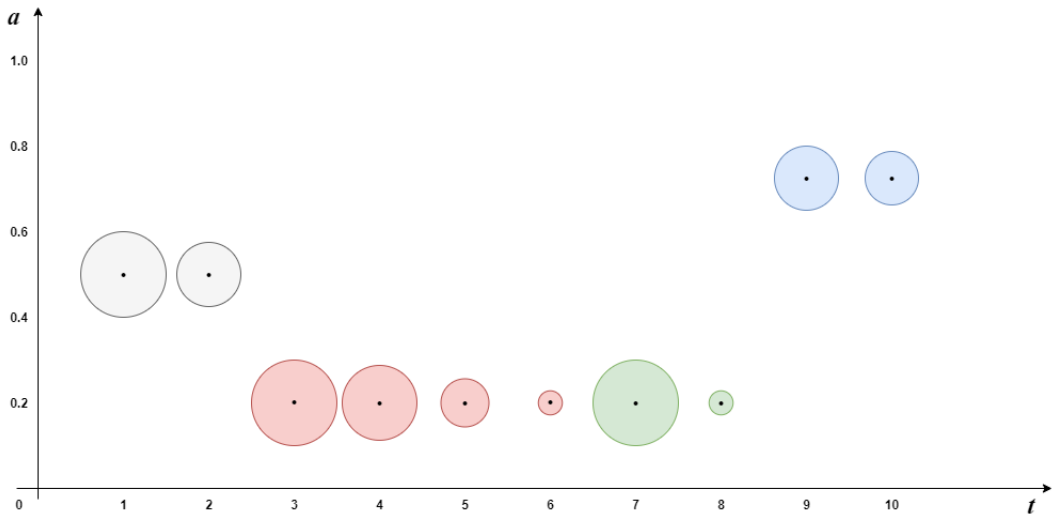


Figure 1.2: Agreements across different periods. *The partisan lean of implemented outcomes is graphed against periods of the game, with the efficiency of an outcome indexed by the radius of the circle circumscribing it.*

efficiency life of a given agreement is the appropriately discounted sum of the efficiencies e of each outcome (e, a) corresponding to the agreement a while it is in place. Since we can interpret the efficiency of an outcome as the collective surplus it yields for both bargaining parties, the efficiency life represents the total surplus an agreement generates to both agents while in place. We make also the following simplifying assumption—

Assumption 2. *Whenever policy is renegotiated, lawmakers are constrained to choose an outcome that is fully efficient. That is, if* $o_t = 0$, *then* $e_t = 1$.

Assumption 2 comports with the examples that motivate the model. For instance, when politicians pass legislation that is meant to capture new technologies that require an updated regulatory framework, it is reasonable to assume that the new environment ought to, in fact, capture all features of that new innovation. Politicians would then contend on the nature of the regulation (e.g., more or less invasive) subject to this natural constraint.

1.2.2 Agreements and Foresight Horizon

I now formalize a key feature of the model.

Definition 2. *We call $T \in \mathbb{N} \cup \{\infty\}$ the **foresight horizon** of a given agent if the agent optimizes at every decision node over exactly the next T agreements.*

Agents with $T < \infty$ exhibit *limited foresight* and curtail at each decision-node the dynamic linkage of agreements resulting from the current outcome to a chain of length T. That is, while agents with *foresight horizon* T recognize that the stream of utility-relevant policies will continue on forever, they plan as if the agreement can only be revised T times. The efficiency of the T^{th} agreement to be instantiated continues to decay without remedy.[10] T thus represents an estimated truncation relevant to the decision-making of the agents, *not* the *actual* horizon of negotiations themselves.

There are several interpretations of this behavior. For instance, it is natural to think that negotiators cannot possibly account for the infinite stream of consequences that flow from a given decision in the current period, and thus limit attention to a finite number of impending agreements. Additionally, parties to negotiations, in practice, only consider those policy revisions that could fall within the span of their own careers, thus again limiting attention to a certain number of downstream re-negotiations. Related notions of limited foresight have been introduced in dynamic, repeated games— for instance, Jehiel (1995), Jehiel (1998), Jehiel (1999), and Wichardt (2010). In online Appendix C, I provide a means of formalizing limited foresight as a natural outcome of bounded rationality and the bargaining environment presented in this Section.

By Definition 1, we can map any sequence of outcomes $\mathbf{p} := \{p_t\}_{t=0}^{\infty}$ onto a corresponding sequence of agreements $\{a^k\}_{k=0}^{\infty}$ through the projection in the second coordinate. For each p_t,

$$p_t \mapsto a^k \text{ s.t. } p_t = (e_t, a^k).$$

[10] This formulation is not necessary for the results derived in this paper. For instance, one could just as easily assume the game ends once the final agreement in the T-length chain is passed. The specification in this paper helps us disentangle the effect of the time horizon and time-discounting on gridlock from that of foresight.

Using again Definition 1 for efficiency life, we can rewrite the utility agent i with foresight T from a sequence of outcomes $\mathbf{p}$ as[11]

$$U_{x_i}^T(\mathbf{p}) = \sum_{t=0}^{\infty} \delta^t u_{x_i}(p_t) = \sum_{t=0}^{\infty} \delta^t e_t \cdot (1 - |x_i - a^k|) = \sum_{k=0}^{T} \delta^{t_k} \eta^{a^k} (1 - |x_i - a^k|), \tag{1.1}$$

We assume that both agents have the same foresight horizon T. We will also be interested in the behavior of the "standard," complete-forecasting agent and how this differs from the finite T case.

1.2.3 Strategies and Equilibrium Concept

For any foresight horizon T exhibited by agents, we will focus throughout on *Markov strategies*, which depend only on $(\hat{p}, \tau)$, where $\tau \in [T] := \{T, T-1, \ldots, 1\}$ is the number of agreements remaining on the foresight horizon. Since strategies are Markov, I will drop the subscript 't' denoting the period from the notation for the status quo and outcome, as it is extraneous.

Definition 3. *A (pure) strategy σ_i^T for agent $i \in \{L, R\}$ when the foresight horizon is T is a tuple of functions $\{(A_i^\tau, p_i^\tau)\}_{\tau=1}^T$, where for any $\tau \in [T]$,*

1. *$A_i^\tau : P \to 2^{[0,1]^2}$ maps the Markov state to the set of outcomes that i accepts at status-quo $\hat{p}$ when i is in the opposition for τ agreements remaining on the horizon*

2. *$p_i^\tau : P \to [0,1]^2$ represents the proposal i makes when the incumbent and there are τ agreements remaining on the horizon.*

We let the components of a given proposal be decomposed as $p_i^\tau(e, \hat{a}) = (1, a_i^\tau(e, \hat{a}))$. Notice that strategies specify the behavior of parties up to T agreements from the current period because the projected equilibrium stream of outcomes is the relevant argument of the objective function for all players. This is in spite of the fact that, given that a T-foresight horizon is used at *each decision node*, the strategies given by A_i^τ, p_i^τ will not be utilized on the actual path of play for any

[11] I use the convention here that if the initial status-quo $\hat{p}_0 = (e, a^0)$ is never enacted, then $\eta_{a^0} = 0$.

$\tau < T$. Instead, these serve as estimates for future play for agents that support the current-period action.[12] The equilibrium concept that will interest us is Markov perfect equilibrium, to which— for simplicity— we will henceforth refer simply as *equilibrium*.

By way of notation, call $p^T(\hat{p}, I, \sigma)$ the outcome enacted in the current period of the game when the state is $(\hat{p}, T)$, the incumbent is I, and agents play strategy $\sigma := (\sigma_L, \sigma_R)$. Let also $a^T(\hat{p}, I, \sigma)$ be the partisan lean of this outcome.

1.3 Determinants of Gridlock

In this section, I analyze equilibrium in the general setting and describe conditions under which gridlock emerges in equilibrium play. First, we internalize the following basic observation.

Lemma 1. *For given equilibrium strategies, define*

$$A^\tau(\hat{p}) := A_L^\tau(\hat{p}) \cap A_R^\tau(\hat{p}).$$

If outcome (e, a) comes into place at $(\hat{p}, \tau)$ for some incumbent on the equilibrium path, then $(e, a) \in A^\tau(\hat{p})$. Moreover, if $A^\tau(\hat{p}) \neq \emptyset$, then some $(1, a) \in A^\tau(\hat{p})$ comes into place in equilibrium for any incumbent.

Lemma 1 characterizes the set of all proposals that would feasibly generate agreement at $(\hat{p}, \tau)$ as the intersection of the sets of acceptable proposals for each of the agents. We now comment on the structure of equilibrium.

Definition 4. *Suppose the foresight horizon of agents is T. We say that an equilibrium of the game is **in cut-off strategies** if $\forall \tau \leq T$, whenever agreement happens on the equilibrium path at $((e, \hat{a}), \tau)$, agreement also occurs at $((e', \hat{a}), \tau)$ for every $e' \leq e$.*

If an equilibrium is in cut-off strategies, it mandates that for any number τ of agreements remaining on the foresight horizon, there exists some cut-off in the efficiency e^* (possibly dependent

[12] This is because the foresight horizon from any node is T, not any $\tau < T$.

on both τ and $\hat{a}$) such that both parties come to an agreement whenever $e \leq e^*$. Our approach will be to show that every equilibrium of the bargaining game is in cut-off strategies as well as to characterize the unique efficiency cut-offs for any equilibrium, given the utility parameters (δ, T) and stochastic decay shock $\tilde{\beta}$.

For ease of exposition, I define the following recursive sequence. Let

$$\begin{aligned} S_0 &:= \frac{1}{1-\delta\overline{\beta}}, \qquad \text{and for } T \geq 1, \\ S_T &:= \mathbb{E}\left[\sum_{t=0}^{\tilde{t}_T - 1} \delta^t \prod_{i=1}^{t} \tilde{\beta}_i\right] + \mathbb{E}\left[\delta^{\tilde{t}_T}\right] S_{T-1}, \end{aligned} \tag{1.2}$$

where $\tilde{t}_T := \min\left\{t : \prod_{i=1}^{t} \tilde{\beta}_i < (1-\delta)S_{T-1}\right\}$. Note that *if* for any agreement $\hat{a}$, agreement is reached if and only if $e \leq (1-\delta)S_{\tau-1}$ with τ agreements remaining on the foresight horizon, then S_{T-1} indexes the expected surplus from the point an agreement is enacted, with foresight horizon T. We show now that this sequence of surpluses is the relevant, self-enforcing one in any equilibrium.

Proposition 1 (Agreement Condition). *For any $T \in \mathbb{N}$ and state $((e, \hat{a}), T)$, every equilibrium of the game is in cut-off strategies with unique threshold $e_T^* = (1-\delta)S_{T-1}$ for agreement to occur.*

The threshold for agreement to occur has a natural interpretation— bargaining agents will only revise the agreement if the status quo efficiency is below the opportunity cost— $(1-\delta)S_{T-1}$— induced by a single period of delay, *given a foresight horizon* T. In particular, since S_{T-1} does not depend on the incumbent selection rule, we have the following corollary:

Corollary 1 (Independent Welfare). *The total surplus S_{T-1} is independent of the mechanism used to select the incumbent in each period.*

Moreover, for each T, since $S_{T-1} < \frac{1}{1-\delta}$, it follows that each threshold for agreement $e_T^* < 1$, for any finite T. Therefore, the event of the status quo efficiency e exceeding the threshold e_T^*, and agreement thus *not* being reached in any equilibrium, occurs in some period generically with positive probability. In this case, we say that *gridlock* occurs. Observe that any cumulative distribution F

over $\tilde{\beta}$ for which $F(x) < 1$ for any $x < 1$ will yield gridlock with positive probability for *every* finite foresight horizon T. This crucial corollary thus follows—

Corollary 2 (Gridlock). *For any $T \in \mathbb{N}$, gridlock occurs in any equilibrium of the game with positive probability for almost all non-atomic shock distributions $\tilde{\beta}$.*

We thus establish general conditions under which gridlock is a feature of any equilibrium. In particular, if gridlock occurs with positive probability in the game, then along the path of equilibrium play, when policy is revised *ad infinitum*, gridlock will occur almost surely. In the next section, we delineate the ways in which δ and T affect equilibrium, particularly, in different ways.

1.3.1 The Incidence of Gridlock

Given the characterization of the efficiency threshold for agreement in Proposition 1, I now examine how the various parameters of the model affect the incidence of gridlock. An important preliminary observation is that the threshold efficiency

$$e_T^* = (1 - \delta)S_{T-1}, \tag{1.3}$$

is precisely the difference in total surplus to bargaining parties from coming to an agreement in the current period as compared against reaching agreement the next period. Since current efficiency e is forgone whenever agreement is reached, the rule $e < e_T^*$ *maximizes aggregate efficiency relative to a foresight horizon T.*

This way of representing the threshold efficiencies gives us a straightforward way of determining the effects the discount factor and foresight horizon of players, respectively, have on the incidence of gridlock. Note that as T increases, the projected future surplus S_{T-1} of both agents, also *increases*; hence, e_T^* increases with T. On the other hand, as the opposition in the game becomes "more patient," she is more inclined to wait for power as the opportunity cost $(1 - \delta)S_{T-1}$ of delay *decreases*. This leads to the following observation.

Proposition 2 (Effects on Gridlock). *In any equilibrium of the game,*

(a) for $T' < T$ and any set of other fixed parameters, $e^*_{T'} < e^*_T$,

(b) for any $T \in \mathbb{N}$, e^*_T is strictly decreasing in δ.

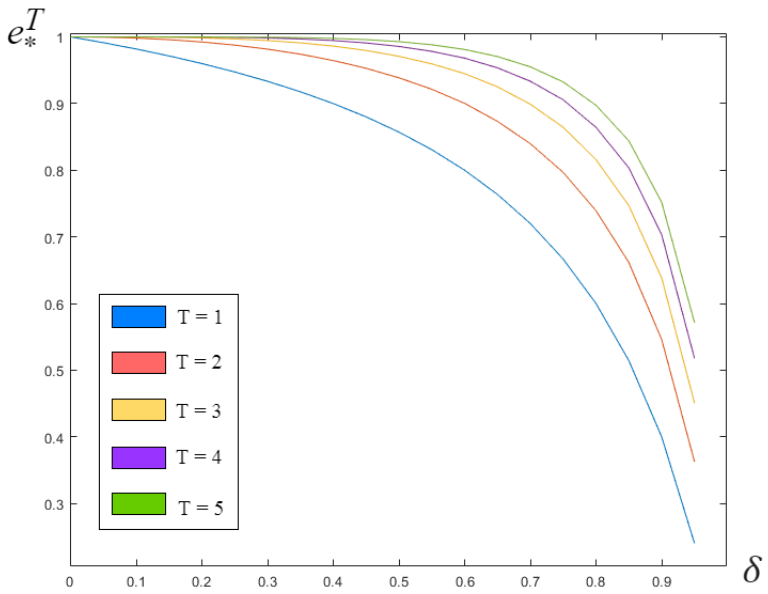


Figure 1.3: As the discount factor δ increases, the threshold for agreement decreases, making compromise more difficult.

Recall that with foresight horizon T, agreement occurs at $\tilde{t}_T$, the minimal period t for which the decreasing sequence $\left(\prod_{i=1}^{t} \tilde{\beta}_i\right)_{t=1}^{\infty}$ falls below threshold level e^*_T. Since neither δ nor T affect this sequence, Proposition 2 shows that lengthening the foresight horizon *decreases* the incidence of gridlock, while increasing the intertemporal discount factor δ *increases* it.

At first glance, δ and T both appear to index a kind of patience. Indeed, δ measures the rate of inter-temporal trade-off, and T measures how many future agreements are incorporated into agents' decision-making in the current period. Thus, small δ and small T both capture the relative importance of the *current* negotiation to agents. How is it then that Proposition 2 reveals that T and δ have *opposite effects* on gridlock?

This can be explained as follows: as T increases, agents incorporate the possibility of further future renegotiation into their utility framework. As a result, the bargaining parties place an increased

priority on refurbishing the efficiency of policy since there are greater losses in surplus from not doing so; simultaneously, when accounting for greater revisions of policy, agents internalize that the partisan lean of the next agreement matters *less* for their long-run utility. Through both these channels, increased foresightedness leads agents to come to agreement more easily. On the other hand, as δ increases, the aggregate efficiency lost by waiting an additional period to come to an agreement is given by $(1-\delta)S_{T-1} - e$. Recall S_{T-1} is the total expected surplus to be split by the agents in the future when the foresight horizon is T. Hence, since δ indexes inter-temporal tradeoffs, large δ makes waiting more desirable relative to arriving at a compromise in the current period.

1.3.2 Infinite Foresight Horizon

With the general result on the structure of equilibrium, along with its implications in hand, we move to a discussion of what the limiting equilibrium behavior entails in the case where $T \to \infty$.

Proposition 3 (**Limiting Equilibrium**). *Let* $\lim_{T\to\infty} e_T^* := e_\infty^*$. *Then,* $e_\infty^* = 1$*; there is no delay on the equilibrium path for any stochastic process* $\tilde{\beta}$.

Recall that each threshold e_T^* maximizes the total surplus to both agents, given a foresight horizon T. If policy were renegotiated *ad infinitum*, then the maximal surplus that could be achieved would be $\frac{1}{1-\delta}$, by coming to an agreement each period. This agreement protocol corresponds to a limiting threshold of $e_\infty^* = 1$. By contrapositive, as encapsulated in Corollary 2, whenever $T < \infty$, the maximal surplus achievable by the two parties is less than $\frac{1}{1-\delta}$, and so given δ, there will always be some $e < 1$ such that it is surplus-maximizing relative to a T-foresight horizon to enact status-quo $(e, \hat{a})$ in the current period rather than renegotiate.

Proposition 3 thus establishes the crucial role that limited foresight plays in generating gridlock. Agents that forecast and value the full stream of policy consequences that flow from the current-period decision would, for any underlying stochastic decay process, always arrive at an immediate agreement. As noted in Section 1.1, there are many reasons— bounded rationality, finite career span, and underlying beliefs— that reify the relevance of limited foresight in political decision-making. I

provide a formal mechanism through which bounded rationality in the form of a "computational budget" leads to limited foresight in online Appendix C.

Importantly, models of policy-making in political science that emerged from the school of *incrementalism* have long emphasized the role of limited foresight— e.g., Lindblom (1959), Gregory (1989), Hayes (2017). We therefore explore an important case of the general model— legislative bargaining with endogenous elections— in the following Section.

1.4 Legislative Bargaining

We now examine a particular instantiation of the general setting: legislative bargaining. Section 1.4.1 presents the specific additions we make to the general model of Section 1.2. In particular, we describe a voting protocol that endogenizes proposal power in the strategies of bargaining agents and allows us to solve for closed-form proposals made by politicians in equilibrium. I assume that there is uncertainty over the position of the median voter, who is independently and identically distributed uniformly within the model's ideological $[0, 1]$-space in each period. In Section 1.4.2, I demonstrate *outcome uniqueness* of equilibrium— conditional on agreement being reached at a given state, the strategies of agents are always unique in any equilibrium. I explicitly solve for the proposals of agents and further describe how the gap between them varies with the foresight horizon T and with δ.

Furthermore, in Section 1.4.3, we study the distribution of policies in the steady state of long-run equilibrium play. I show that the ideological lean of policies converges to an invariant distribution symmetric about the expected median voter, $\frac{1}{2}$— an attenuated long-run version of the Median Voter Theorem. I discuss how the foresight horizon T modulates the variance of this invariant distribution and the implications of this relationship. Finally, Section 1.4.4 extends many of the results of the baseline legislative bargaining environment to a more general setting where we allow the median voter to be distributed according to a *family* of distributions, indexed by γ, which governs the variance in the electorate's median ideological position in any period. I show that as the variance

decreases (γ increases), party's proposals may move in the *same* ideological direction, with the midpoint of party's proposals always moving closer to the expected median ideology, $\frac{1}{2}$. The effect of uncertainty over the median voter's position is thus different from the standard results from the electoral spatial competition literature— a point I explore.

1.4.1 The Setting

I now build on the general model outlined in Section 1.2, elaborating on the specific features of the legislative bargaining setting. In effect, we will embed a particular rule for selecting political incumbents within the general framework.

The electorate is modeled as a continuum of voters, with each voter j identified with her current favorite policy $x_j \in [0, 1]$. For any given set M of outcomes from which a voter chooses, let

$$||M|| := \max_{p,p' \in M} |a_p - a_{p'}|$$

be the maximum ideological distance between any two policies within the electorate's menu M. A voter with ideal point x_j has a utility from outcome $p = (e_p, a_p)$ from within a menu M given by

$$v_{x_j}(p|M) = -\frac{e_p \cdot (x_j - a_p)^2}{||M||}. \tag{1.4}$$

The numerator of $v_{x_j}(p|M)$ is the efficiency of policy, e_p, multiplied by the usual quadratic utility that voters often exhibit in models of spatial competition. We normalize this standard utility by dividing by the maximum ideological distance between any two outcomes (alternatives) the voter faces from within her "electoral" menu, M. Hence, voters measure the distance from their own ideal point to the ideological lean of a given policy p in *relative terms*, with reference to the total span of platforms politicians offer at any given point of time.[13]

[13] Alternate utility frameworks also yield the results of Section 4. For instance, one could also impose that voters have preferences analogous to the parties'— $v_{x_j}(p) = e_p \cdot (1 - |x_j - a_p|)$ — but exhibit *single-period myopia* so that they vote for the candidate that induces the preferred current-period outcome.

As in the case of the parties, I assume that the stochastic decay shock $\tilde{\beta}_t$ is immediately observed by all players so that the inherited status quo $\hat{p}_t$ is common knowledge. Once $\hat{p}_t$ materializes, an election is held to determine the composition of government at time t. Voters have perfect information over the policy preferences of political parties, and akin to politicians, also have foresight horizon T. Crucially, politicians lack the ability to credibly commit to a platform, and instead, each voter selects the party that, as the current-period incumbent, would induce the more preferred stream of outcomes in expectation over a horizon of T agreements in equilibrium. That is, voter j strictly prefers party R as the current incumbent if and only if

$$\sum_{t=0}^{\infty} \delta^t \mathbb{E}[v_{x_j}(\tilde{q}_t^T(R)|M_t)] > \sum_{t=0}^{\infty} \delta^t \mathbb{E}[v_{x_j}(\tilde{q}_t^T(L)|M_t)],$$

where $\tilde{q}^T(I) := \{\tilde{q}_t^T(I)\}_{t=0}^{\infty}$ is the *projected*, infinite sequence of stochastic outcomes that results in equilibrium from electing $I \in \{L, R\}$ the current incumbent, when all agents have foresight horizon T, and M_t is the period-t menu of policies. Each voter elects her preferred party, randomizing in the case of indifference. The party that garners a majority of votes takes office as the incumbent; the loser is the opposition. The game then proceeds as before, and is repeated *ad infinitum*.

Hence, the electorate plays a key role in determining *endogenously* the dynamic allocation of proposal power between the two parties, and thus also influences their incentives. I now make one further assumption to stipulate precisely the electoral environment. We assume[14] that there is uncertainty over the identity of the median voter m_t at each time t. In particular,

Assumption 3. *The median voter $m_t \in [0, 1]$ in the population is distributed uniformly on [0,1]. The median m_t realizes at the beginning of period t and is unobservable to all agents.*

This is a baseline assumption made for tractability that we will relax in Section 1.4.4. Finally, as it is convenient notation throughout the proofs and exposition that remains, we let $\tilde{\eta}^\tau$ be the efficiency life of an agreement when there are $\tau \in [T]$ agreements remaining on the foresight horizon (as in the appendix). We denote by I_τ the incumbent when agreement is reached with τ

[14] As in many standard models of spatial competition in political economy— e.g., Wittman (1983), Poole (2005)

agreements remaining on the foresight horizon. We define

$$W_i^T(a) := u_{x_i}(1,a) \cdot \mathbb{E}[\tilde{\eta}^{T-1}] + \mathbb{E}[\delta^{\tilde{t}_{T-1}} u_{x_i}(1, a^{T-1}((\tilde{e}^{T-1}, a), I_{T-1}, \sigma))) \cdot \tilde{\eta}^{T-2}] + \dots$$

as the total utility to political party $i \in \{L, R\}$ from implementing policy $(1, a)$ in the current period when her foresight horizon is T and downstream strategies are played according to σ. The following lemma will prove useful in characterizing equilibrium—

Lemma 2. *Consider σ such that in any state of the world $\hat{p}$ at which agreement occurs, $a_R^T(\hat{p}) > a_L^T(\hat{p})$. Then, the pivotal voter $\mu^T(e, \hat{a})$ indifferent between R and L at any status quo $(e, \hat{a})$ at which agreement occurs, is*

$$\mu^T(e, \hat{a}) = \frac{W_R^T(a_R^T(e, \hat{a})) + W_R^T(a_L^T(e, \hat{a}))}{2S_{T-1}}.$$

With the particulars of the legislative bargaining setting now outlined, we can delve into the characterization of equilibrium within this model.

1.4.2 Equilibrium Characterization

In Proposition 2, we showed that the states in which agreement is reached are the same in any equilibrium, characterized by an efficiency cut-off for agreement that depends only on the foresight horizon T. The legislative bargaining environment allows us to further demonstrate that in any equilibrium, whenever agreement is reached, the specific proposal passed is unique, conditional on the identity of the incumbent and the status quo policy. This notion of uniqueness within equilibria is formalized here.

Definition 5. *For any foresight horizon T, we say that the game has an **outcome-unique** equilibrium if for each $\tau \leq T$, whenever σ and σ' are strategy profiles played in any two equilibria, $p^T(\hat{p}, I, \sigma) = p^T(\hat{p}, I, \sigma')$ for any $\hat{p} \in P$ and $I \in \{R, L\}$.*

Whenever $A^T(\hat{p}) = \emptyset$ in equilibrium, it follows that *any* proposal not in A^T_{-I} made by the incumbent I will support such an equilibrium. Hence, uniqueness of equilibrium strategies in general is infeasible, so that outcome-uniqueness is the *relevant* and strongest notion of uniqueness obtainable. In the main proposition of Section 1.4 below, I demonstrate that there is an outcome-unique equilibrium of the game and solve for the (unique) proposal functions of both parties conditional on agreement.

Proposition 4 (**Equilibrium Characterization**). *For any $T \in \mathbb{N}$, there is an almost-everywhere outcome-unique equilibrium of the game that implements unique proposals $p_i^T(\hat{p})$ with $e_i^T(\hat{p}) = 1$ and $a_i^T(e, \hat{a})$ linear and increasing in $\hat{a}$ for $i \in \{R, L\}$ for any $\hat{p} = (e, \hat{a})$ at which agreement occurs on the equilibrium path. We write for each i,*

$$a_i^T(e, \hat{a}) = C^T(e)\hat{a} + B_i^T(e),$$

where $C^T(e) \in (0, 1)$ for $e > 0$ and $C^T(0) = 0$.

By Proposition 1, for arbitrary T, any equilibrium must be in cut-off strategies with threshold $e_T^* = (1 - \delta)S_{T-1}$, as defined by (1.2). Whenever $e > e_T^*$, there is no new agreement implemented; this lack of agreement implies that for any equilibrium, the outcome for any $e > e_T^*$ is given by the status quo policy. Thus, proving the proposition is equivalent to showing uniqueness of the policy proposals and solving for their linear form, conditional on agreement.

Our approach in doing so will be to use induction on the foresight horizon T. Here, I provide the tenor of the analysis by solving the base case $T = 1$. Consider $\hat{p} = (e, \hat{a})$ such that $e \leq e_1^*$. Using (1.13), the problem of incumbent $I \in \{R, L\}$ is given by

$$\max_a \left\{ u_{x_I}(1, a) \cdot \frac{1}{1 - \delta\overline{\beta}} \right\} \quad \text{subject to} \quad u_{x_{-I}}(1, a) \cdot \frac{1}{1 - \delta\overline{\beta}} \geq e\hat{a} + \delta\mathbb{E}\left[V_i^1(\tilde{\beta}_1 e, \hat{a}|\sigma) \right],$$

which can be re-expressed as

$$\max_a \left\{ \frac{a}{1-\delta\overline{\beta}} \right\} \quad \text{subject to} \quad \frac{1-a}{1-\delta\overline{\beta}} \geq e(1-\hat{a}) + \delta\mathbb{E}\left[V_L^1(\tilde{\beta}_1 e, \hat{a}|\sigma) \right] \tag{1.5}$$

$$\max_a \left\{ \frac{1-a}{1-\delta\overline{\beta}} \right\} \quad \text{subject to} \quad \frac{a}{1-\delta\overline{\beta}} \geq e\hat{a} + \delta\mathbb{E}\left[V_R^1(\tilde{\beta}_1 e, \hat{a}|\sigma) \right] \tag{1.6}$$

for R and L, respectively. From (1.5) and (1.6), it is immediate that the R-incumbent will implement the policy with the maximal partisan lean acceptable to L and the L-incumbent the policy with the minimal partisan lean acceptable to R.[15] Hence, the proposal functions of the two agents satisfy

$$\frac{1-a_R^1(e,\hat{a})}{1-\delta\overline{\beta}} = e(1-\hat{a}) + \delta\mathbb{E}\left[V_L^1(\tilde{\beta}_1 e, \hat{a}|\sigma) \right] \quad \text{and} \quad \frac{a_L^1(e,\hat{a})}{1-\delta\overline{\beta}} = e\hat{a} + \delta\mathbb{E}\left[V_R^1(\tilde{\beta}_1 e, \hat{a}|\sigma) \right]. \tag{1.7}$$

Rearranging and summing the two equations in (1.7) gives us

$$a_R^1(e,\hat{a}) - a_L^1(e,\hat{a}) = 1 - \delta - e(1-\delta\overline{\beta}) := \Delta^1(e). \tag{1.8}$$

That is, the two proposal functions differ by an amount $\Delta^1(e)$ that depends only on the status quo efficiency. We solve (1.7) for the L-incumbent's proposal function. Using the fact that agreement must occur at any $e \leq e_1^*$ and letting π_L be the probability that L wins the next-period election,

$$\frac{1}{1-\delta\overline{\beta}} \cdot a_L^1(e,\hat{a}) = e\hat{a} + \frac{\delta}{1-\delta\overline{\beta}}\mathbb{E}\left[\pi_L \cdot a_L^1(\tilde{\beta}_1 e, \hat{a}) + (1-\pi_L) a_R^1(\tilde{\beta}_1 e, \hat{a}) \right].$$

Moreover, since the hypothesis of Lemma 2 is satisfied, each voter j with

$$x_j < \frac{a_L^1(\tilde{\beta}_1 e, \hat{a}) + a_R^1(\tilde{\beta}_1 e, \hat{a})}{2}$$

will vote for L. Substituting in for π_L and further using (1.8), we obtain

[15] In the general, this amounts to showing that R's utility is increasing in the partisan lean of policy she implement. This is trivial for the $T = 1$ case and requires a bit of work in the general case $T \geq 2$, which is provided in the Appendix.

$$a_L^1(e,\hat{a}) = e\hat{a} + \delta\mathbb{E}\left[\left(1 - \frac{\Delta^1(\tilde{\beta}_1 e)}{2}\right)\Delta^1(\tilde{\beta}_1 e)\right] + \delta\mathbb{E}\left[(1-\Delta^1(\tilde{\beta}_1 e))a_L^1(\tilde{\beta}_1 e,\hat{a})\right]. \tag{1.9}$$

The right-hand side of (1.9) has an $a_L^1(\tilde{\beta}_1 e,\hat{a})$ term. By continuing to re-apply the expansion identity (1.9) on the right-hand side recursively, and noting that because each successive term $a_L^1(\prod_{i=1}^k \tilde{\beta}_i \cdot e,\hat{a})$ will have a δ^k coefficient, the recursion process will converge to a solution for $a_L^1(e,\hat{a})$—

$$a_L^1(e,\hat{a}) = \underbrace{e(1-\delta\overline{\beta})\left(\sum_{k=0}^{\infty}\delta^k\mathbb{E}\left[\prod_{i=1}^{k}\left(\delta + \tilde{\beta}_1\cdot\ldots\cdot\tilde{\beta}_i e\right)\tilde{\beta}_i\right]\right)}_{C^1(e)}\cdot\hat{a} + \underbrace{\frac{1}{2}\left(\sum_{k=1}^{\infty}\delta^k\mathbb{E}\left[\left(1-\left(\delta+\tilde{\beta}_1\cdot\ldots\cdot\tilde{\beta}_k e\right)^2\right)\prod_{i=1}^{k-1}\left(\delta+\prod_{j=1}^{i}\tilde{\beta}_j\cdot e\right)\right]\right)}_{B_L^1(e)}. \tag{1.10}$$

Clearly, (1.10) is linear and increasing in $\hat{a}$ since $C^1(e) \geq 0$ (and > 0 for any $e > 0$). Since $a_R^1(e,\hat{a}) = a_L^1(e,\hat{a}) + \Delta^1(e)$, a_R^1 is also linear and increasing in $\hat{a}$.

This analysis proves Proposition 4 in the case where $T = 1$. An induction argument along with some supplementary technical work (see Lemmas 4 - 1.B.1) generalizes the result to $T > 1$.

Proposition 4 establishes outcome-uniqueness of equilibrium, except for the measure zero event when $e = e_T^*$, along with the linearity of *actionable* policy proposals in the status-quo agreement. In particular, proposals of *both* parties are positively correlated with the status quo partisan lean $\hat{a}$. Because this holds for every $\tau \leq T$, the entire stream of projected future policies is positively correlated with the direction of new agreement in the current period— an attenuated persistence in policy. The following result naturally follows—

Corollary 3. *When $e \leq e_T^*$, $\mathbb{P}$[R wins] is decreasing in $\hat{a}$ and $\mathbb{E}[a^T(e,\hat{a})]$ is increasing in $\hat{a}$.*

That is, the passage of extremal policies negatively impacts the corresponding party's chances in future elections, but shifts future policy in the direction of the extremal policy. The effect is entirely *prospective*: a policy enacted at the ideological extremes induces a favorable disagreement point for the corresponding party. In order to induce agreement, then, that party is less willing to compromise on the partisan lean of policy. This in turn shifts the agreement set in the direction of the new status-quo, accounting for why future policies correspondingly shift, in expectation. On the other hand, this means that the *opposite* party will propose the more moderate policy and hence has a greater chance of being elected in the future.

Yet another feature of equilibrium is the ideological distance between the equilibrium actionable proposals made by potential incumbents. For general T, we write this distance as $\Delta^T(e)$, noting that because of the linearity of proposals in $\hat{a}$, this distance will not depend on the status quo partisan lean, as in (1.8). Using the properties of equilibrium gleaned from Proposition 4, we show that this gap increases with foresight decreases in the inter-temporal discount rate δ.

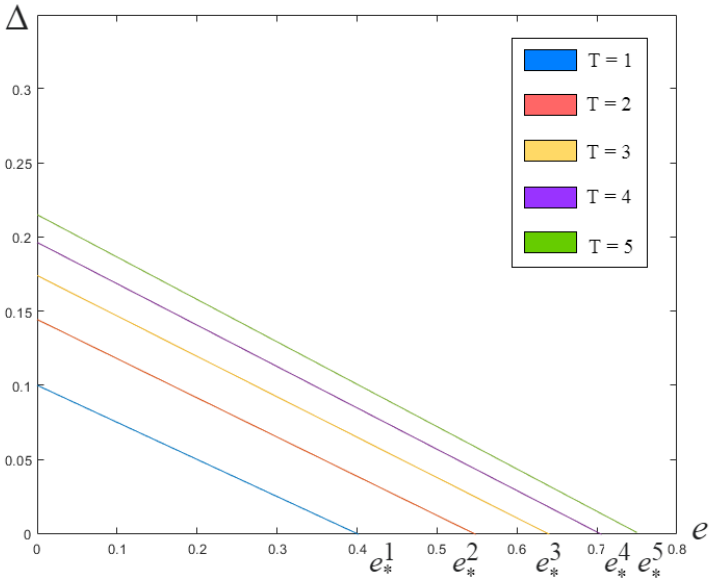


Figure 1.4: Comparative Statics on T. *For any* $T > T'$, $e_*T' < e_T^*$ *and* $\Delta^T(e) > \Delta^{T'}(e)$ *for* e *at which agreement occurs*

Proposition 5. *Consider $T' < T$. Whenever $e \leq e^*_{T'} < e^*_T$, $\Delta^T(e) > \Delta^{T'}(e)$, with strict inequality whenever $e \neq 0$. Moreover, $\Delta^T(e)$ is strictly decreasing in δ.*

Thus, the effects the foresight horizon and the inter-temporal discount factor have on the distance between parties' proposals are the opposite as their effects on gridlock. Importantly, this should not be confused with the notion that gridlock and polarization are strategic substitutes. Rather, when the extent of gridlock is greater (with small T, for example), it is often unlikely that any legislation will pass imminently. Yet, conditional on new policy being enacted in such a scenario, it tends to fall within a narrow, centrist window since the set of mutually acceptable policies is smaller. Thus, the distance between the leftmost and rightmost proposals that could be approved by both parties under the presence of greater gridlock tends to be small.

Having solved for equilibrium in the legislative bargaining setting, we now turn our attention to the long-run properties of equilibrium play.

1.4.3 Policy Dynamics and Incrementalism

As discussed in Section 1.2, agents with foresight horizon T behave at every node of the game as though their foresight horizon *is* T, irrespective of past agreements hatched by agents. We can then ask what the long run distribution of policies passed in a legislature with policy-makers of foresight horizon T will look like.

To begin answering this question, we must first show that there is indeed some convergent policy distribution for every foresight horizon. Consider first two different policies $x, y \in [0, 1]$. Assume without loss of generality that $x < y$ and let $\pi_i^T(z, e)$ be the probability that i wins the election when policy z is to be revised at efficiency level $e < e^*_T$. We let H denote the cumulative distribution of

$\tilde{e}^T$. We then have

$$\begin{aligned}
\mathbb{E}\left[\left|a^T(\tilde{e}^T,x)-a^T(\tilde{e}^T,y)\right|\right] &= \int_{[0,e_T^*]}\Big\{\left|a_L^T(e^T,x)-a_L^T(e^T,y)\right|\cdot\pi_L^T(x,e^T)+\left|a_R^T(e^T,x)-a_R^T(e,y)\right|\cdot\pi_R^T(y,e^T)\\
&\qquad +\left|a_R^T(e^T,x)-a_L^T(e^T,y)\right|\cdot(1-\pi_L^T(x,e^T)-\pi_R^T(y,e^T))\Big\}\,dH(e^T)\\
&= \int_{[0,e_T^*]}\Big\{C^T(e^T)\cdot|x-y|\cdot\left(\pi_L^T(x,e^T)+\pi_R^T(y,e^T)\right)\\
&\qquad +\left|a_R^T(e^T,x)-a_L^T(e^T,y)\right|\cdot(1-\pi_L^T(x,e^T)-\pi_R^T(y,e^T))\Big\}\,dH(e^T)\\
&\leq \mathbb{E}\left[C^T(\tilde{e}^T)(1+\Delta^T(\tilde{e}^T))\right]\cdot|x-y|,\ \text{by using the expressions for } \pi_i^T(z,e)\\
&\leq \left(1-\mathbb{E}[B_R^T(\tilde{e}^T)^2]\right)\cdot|x-y|
\end{aligned}$$

Thus, the policy operator acts as a contraction in expectation; over time, the expected distance between policy outcomes that stem from x and y grows closer by factor less than $\alpha := 1 - \mathbb{E}[B_R^T(\tilde{e}^T)^2] < 1$. Notationally, let $\left(a^T\right)^N$ be the application of the stochastic policy operator N times. We then have, through straightforward application of the above inequality,

$$\lim_{N\to\infty}\mathbb{E}\left[\left|\left(a^T\right)^N(e^T,x)-\left(a^T\right)^N(e^T,y)\right|\right]=0.$$

This gives us a strong *no legacy* condition; the effect of the initial policy— x or y— on the distribution of policies after N future agreements are passed, vanishes as N grows large. The same principle can, in fact, be used to show that for any T, there is an invariant distribution over policies in the long-run of equilibrium play. Moreover, that distribution is symmetric about the mean, $\frac{1}{2}$, of the median voter distribution.

Definition 6. *Suppose F and G are cumulative distribution functions over a complete metric space Z. Let $d(F,G) := \int_Z |F(z)-G(z)|\,dz$ be the standard $\mathcal{L}^1$-norm on the space of integrable functions on Z. We say that a sequence of distributions $\{F_n\}_{n=1}^{\infty}$ converges to F^* if $\lim_{n\to\infty} d(F_n,F^*)=0$.*

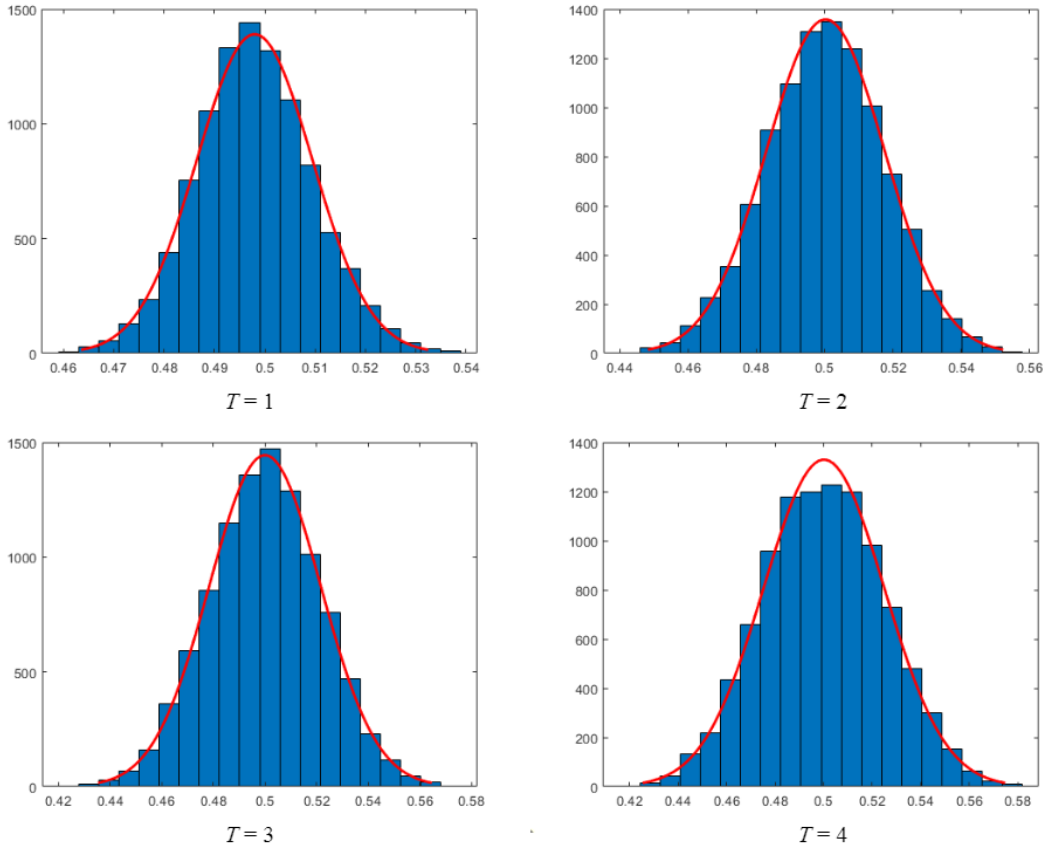


Figure 1.5: The long-run distribution of policy for $T = 1$, 2, 3, and 4. Here, I use β distributed according to a truncated Beta-distribution with $\gamma = 1$, $\beta = 5$.

Proposition 6 (Long-Run Policy Dynamics). *Let $(L_T)^N(F)$ be the distribution of policies after N agreements have been reached between agents with foresight horizon T, when the initial agreement is drawn from distribution F. Then for any F, as $N \to \infty$, $(L_T)^N(F)$ converges to G^* symmetric about $\frac{1}{2}$.*

We thus obtain an attenuated version of the Median Voter Theorem, which asserts that in the long-run, the policy distribution converges in expectation to the expected type of the median voter. We thus retain a key feature of many models in the dynamic electoral accountability literature (See Section 1.1), in which the pressures of catering to the median voter help direct the incumbent towards their ideal point. In our model, because agents cannot credibly commit to a platform, these electoral considerations enter the analysis only through their effect on agents' outside option from vetoing an offer. That is, once a status quo agreement is ideologically close to $\frac{1}{2}$, the value from the

outside option of vetoing a proposal is similar for both potential opposition parties, R and L. The difference in their leverage is minimal. Hence, parties' proposals in this case are near-symmetric about $\frac{1}{2}$. This helps further reinforce the same effect dynamically. Note that even were the position of the median voter fixed through time, there would still be variation in the long-run proposals of players through the realized efficiency $\tilde{e}^T$. This generates a *non-trivial* distribution of policies in the long run.

The variance of the invariant distribution is also of particular interest, as it represents the average distance of the partisan lean of policy from the center of the policy space, $\frac{1}{2}$, in the long-run. This gives us a sense of how "extreme" policies are expected to be in steady-state. Figure 1.5 illustrates that as T increases, the variance of the long-run policy distribution also increases (here, for a Beta-distributed shock). This is a direct consequence of the fact that $\Delta^T(e)$ is increasing in T for a fixed e. Hence, so long as the distribution over "stopping" efficiency-levels $\tilde{e}^T$ at which renegotiation occurs is sufficiently smooth, this observation naturally follows.

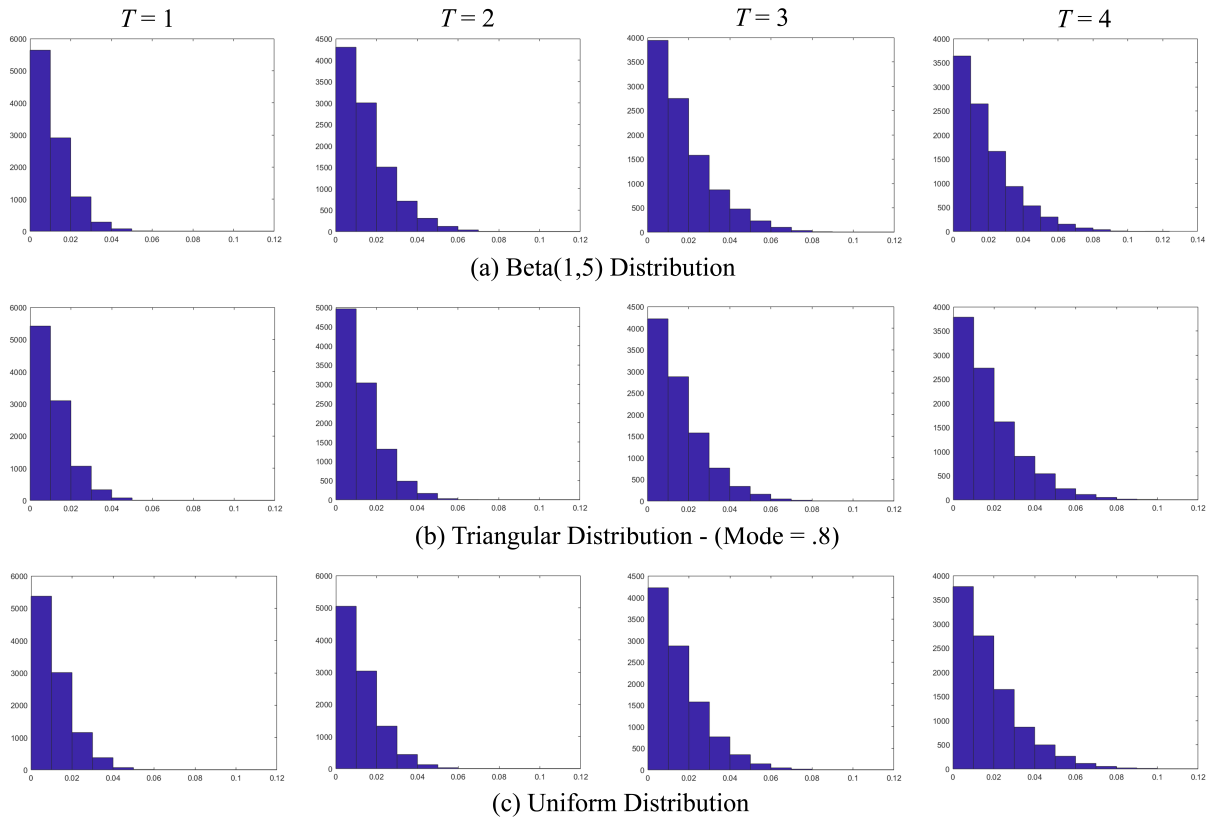
Shock Distribution	$T=1$	$T=2$	$T=3$	$T=4$
$\beta(1,5)$	0.0108	0.0148	0.0173	0.0196
Triangular (mode 0.8)	0.0111	0.0124	0.0155	0.0181
Uniform $[0,1]$	0.0114	0.0121	0.0154	0.0180

Table 1.1: Simulated Mean of Long-Run Policy Increments for Various Shock Distributions, $T \in \{1,2,3,4\}$

Shock Distribution	$T=1$	$T=2$	$T=3$	$T=4$
$\beta(1,5)$	7.53×10^{-5}	1.52×10^{-4}	2.18×10^{-4}	2.92×10^{-4}
Triangular (mode 0.8)	7.91×10^{-5}	1.02×10^{-4}	1.73×10^{-4}	2.43×10^{-4}
Uniform $[0,1]$	8.14×10^{-5}	9.20×10^{-5}	1.66×10^{-4}	2.28×10^{-4}

Table 1.2: Simulated Variance of Long-Run Policy Increments for Various Shock Distributions, $T \in \{1,2,3,4\}$

On the other hand, this logic gives rise to another, even more important realization. As T increases, the absolute difference in the partisan lean from one agreement to the next, the *policy increment*, also increases. This is shown for a variety of different distributions for the decay shock $\tilde{\beta}$ in Figure 1.6 and Tables 1.1 and 1.2. Hence, as agents grow less foresighted, policy-making in



(a) Beta(1,5) Distribution

(b) Triangular Distribution - (Mode = .8)

(c) Uniform Distribution

Figure 1.6: As T increases, both the mean and variance of the policy increments increase in the long-run of equilibrium play.

the long run exhibits more *incrementalism*— policy increments tend to be smaller and more subtle. The observation that swings in policy from one regime to the next tend to be muted in scope was popularized by political scientist Charles Lindblöm whose theory of incrementalism provided both a descriptive as well as normative framework for political decision-making.

Lindblöm's theory involved a heavy critique of models of legislative decision-making that incorporated full rationality. Instead, Lindblom (1959) argues, "[the decision-maker] would not ordinarily find a body of theory precise enough to carry him through a comparison of [policies'] respective consequences. Instead, he would have to rely on the record of past experience with small policy steps." Thus, decision-makers adopt more familiar or predictable policy proposals, given their inability to project the consequences of policies that are substantially different from the current one, and experiment serially rather than engage in radical experimentation. Hayes (2017) similarly encapsulates Lindblöm's central assumption—

> Forecasting consequences accurately requires a very high level of understanding of cause-and-effect relationships among pertinent variables that is unlikely to be attained for most issues. ... Policymakers lack reliable estimates of the consequences associated with each alternative.

The notion of limited foresight in this paper therefore refashions Lindblöm's central assumption in an explicitly *dynamic* context. In this model, agents are unable to parse the consequences a given piece of legislation will have— not on various different interest groups, sectors of the economy, etc., as in Lindblom (1959)— but rather on future negotiations and enactments of policy. Strikingly, then, the above simulations demonstrate that our two-party, limited foresight bargaining game endogenizes as a feature of equilibrium Lindblöm's influential observation.

1.4.4 Uncertainty About the Median Voter

One remaining question is how the *variance* in the underlying distribution of the median voter's ideal policy affects equilibrium. A partial answer is immediately furnished by Proposition 1. Here, the incidence of gridlock is shown to depend only on primitives of the model— the distribution of $\tilde{\beta}$, the foresight horizon T, and the discount factor δ— and not on the rule used to determine incumbency. However, parties' equilibrium proposals do, in fact, shift as uncertainty over the ideological type of the median voter changes. In this section, I show that as this uncertainty decreases, the midpoint of parties' proposals grows closer to the expected median ideal policy, $\frac{1}{2}$, in the electorate. However, rather than proposals *converging* when they lie on either side of $\frac{1}{2}$, as in the spatial competition literature, they move in the *same* direction.

For ease of analysis, we focus on the following family of distributions—

Definition 7. *For $\gamma \in [-4, 4]$, we say the distribution of a random variable X is γ-triangular if the probability density function g_γ of X is given by*

$$g_\gamma(x) = \begin{cases} 1 + \gamma\left(x - \frac{1}{4}\right) & \text{if } x \in \left[0, \frac{1}{2}\right] \\ 1 - \gamma\left(x - \frac{3}{4}\right) & \text{if } x \in \left(\frac{1}{2}, 1\right] \\ 0 & \text{if } |x| > 1 \end{cases}.$$

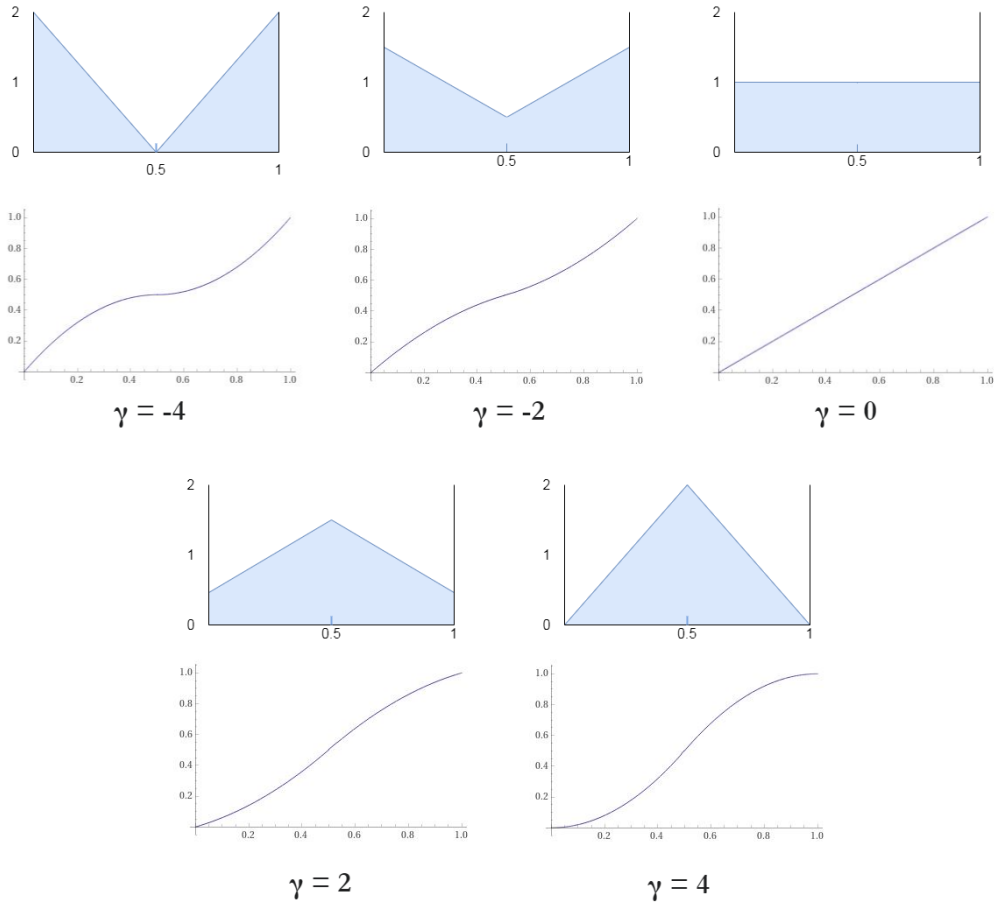


Figure 1.7: The probability distribution functions (top) and cumulative distribution functions (bottom) for alpha-triangularly distributed variables, for each of $\gamma = -4, -2, 0, 2$, and 4.

I call the associated cumulative distribution G_γ. Figure 1.7 shows the probability and cumulative density functions of the $\gamma-$ triangular distribution for a variety of values for γ. Trivially, $\gamma = 0$ gives a uniform density, and thus the model examined in the previous portions of this Section is a special case of the one we now analyze. In particular, we relax Assumption 3 to the following—

Assumption 3'. *In every period t, the median voter $m_t \in [0,1]$ in the population follows an γ-triangular distribution for some $\gamma \in [-4,4]$.*

For clarity, I index with γ each of the terms that depends on this distributional parameter. For instance, the partisan lean of the proposals made by incumbent party I is written as $a_I^T(e,\hat{a}|\gamma)$. The following result describes how proposals vary with the partisan lean of the status quo and γ in the more general legislative bargaining model.

Proposition 7 (**The Effect of Uncertainty Over the Median Voter**). *Each of the following holds in the unique equilibrium of the legislative bargaining model under Assumptions 1, 2, and 3'.*

(i) For each $I \in \{L,R\}$, and any T, a_I^T is increasing in $\hat{a}$ (and strictly increasing for $e > 0$).

(ii) For any T and $I \in \{R,L\}$, when $\hat{a} < \frac{1}{2}$, $W_R^T(a_I^T(e,\hat{a}|\gamma)$ and $a_I^T(e,\hat{a}|\gamma)$ are strictly increasing in γ; when $\hat{a} > \frac{1}{2}$, $W_R^T(a_I^T(e,\hat{a}|\gamma)$ and $a_I^T(e,\hat{a}|\gamma)$ are strictly decreasing in γ. Thus, the midpoint of proposals moves toward (away from) $\frac{1}{2}$ as γ increases (decreases).

(iii) For any T, $W_R^{T+1}(a)$ is strictly increasing in a.

Parts (i) and (iii) of Proposition 7 are generalizations of features of equilibrium in the case $\gamma = 0$ (Proposition 4). For (i), as before, when $\hat{a}$ increases, the leverage of the R-party increases. Since her outside option from disagreement is more favorable, she can both demand larger concessions from the L-incumbent as well as force the L-opposition to accept a more partisan proposal. Statement (iii) directly follows from (i). Because both parties' proposals are increasing in $\hat{a}$, for any T, implementing a more right-leaning proposal in the present will increase the expected partisan lean of all future proposals, thus increasing R's long-run utility from implementing such a policy.

Statement (ii) of Proposition 7, on the other hand, is both novel and subtler, and thus behooves some explanation. To fix ideas, suppose $\hat{a} < \frac{1}{2}$ so that the L party has more leverage and thus always proposes a policy that is *more* ideologically partisan than R. Consider $\gamma' > \gamma$. Under γ', the median voter distribution is more concentrated around $\frac{1}{2}$. If the proposals of the politicians *were to remain fixed*, then L's electoral prospects would diminish and those of R increase (since

L's proposal is more ideologically partisan). Hence, when L is the current incumbent, R's value from the outside option of rejecting L's offer *increases* since R is more likely than before to be the incumbent after the next election. Thus, L must *moderate* her own proposal, moving to some $a_L^T(e, \hat{a}|\gamma') > a_L^T(e, \hat{a}|\gamma)$, in order to induce acceptance from R. On the other hand, the situation is exactly reversed for an R-incumbent, who can extract further concessions from an L-opposition than before.

While the midpoint of politicians' proposals moves closer to $\frac{1}{2}$, in general *both* proposals do not. This distinguishes Statement (ii) from a standard result in spatial electoral competition: as uncertainty over the median voter dissipates, politicians' proposals converge. Our model's dynamics are distinguished from this result since politicians cannot credibly commit to platforms; voters infer how they *would* behave were they to become the current-period incumbent. Thus, altering the voter distribution does *not* induce strategic re-positioning *in order to improve electoral prospects*. Incumbents simply respond to changes *ceteris paribus* in the value of their counterpart's outside option of rejection and adjust their proposals accordingly.

1.5 Conclusion

1.5.1 Extensions

Generalizations of Foresight

The notion of limited foresight studied in this paper thus far places a decision weight $\lambda_t = 1$ on all agreements indexed by $t \leq T$ on the foresight horizon, and $\lambda_t = 0$ for all $t > T$. A generalization of this concept and further research questions that naturally arise are briefly discussed here.

Definition 8. *Let $\lambda \in \mathbb{R}_+^\infty$ be such that $\lambda_{t+1} \leq \lambda_t$ for all $t \in \mathbb{N}$. Recall the coordinate projection mapping $p \mapsto a$ if $p = (e, a)$ for some e. We say that a policy maker exhibits λ-* ***foresight*** *if her*

utility from a stream of outcomes $\boldsymbol{p} = \{p_t\}_{t=0}^{\infty}$ *is given by*

$$U_{x_i}^{\lambda}(\boldsymbol{p}) = \sum_{k=0}^{\infty} \lambda_k \delta^{t_k} \eta_{a^k} |x_i - a^k|. \tag{1.11}$$

Note that λ_k indexes the relative importance of the k^{th} agreement struck in the future modulo exponential discounting, so that *limited foresight* is still captured by this model. Note that all the motivating interpretations of limited foresight continue to hold; for instance, λ_k can be understood to mean the *probabilistic belief* that k revisions of policy will be made or the relative *precision* with which a decision-maker can make computational estimates over policy consequences k agreements downstream from present negotiations.

Example 1 (**Hyperbolic Discounting**). *By hyperbolic discounting, we mean such discounting across agreements, not intertemporally in the direct sense of Laibson (1997). Consider the following sequence of* λ^T- *vectors, where for each* λ^T*,*

$$\lambda_k^T = \begin{cases} 1 & \text{if } k = 0 \\ \lambda & \text{if } k \in 1, 2, \ldots, T \\ 0 & \text{if } k > T \end{cases}$$

λ^T thus places weight 1 on the initial status-quo, $\lambda \in (0, 1)$ on each agreement up to a horizon of T and then disregards all other agreements. Note that in this context, hyperbolic discounting de-values all agreements by factor of λ against the status-quo ("status-quo bias"). For ease of notation, I index all previously defined objects from Section 1.2 with the superscript λ^T when considering equilibrium behavior under λ^T- foresight to distinguish them from the limited foresight objects. The agreement characterization (Proposition 1) can be trivially adapted to this setting. In particular, note that for each $T \geq 1$, for any $\tau \in \{1, \ldots, T-1\}$,

$$p_i^{\lambda^T}(\hat{p}, \tau) = p_i^{\tau}(\hat{p}), \qquad A_i^{\lambda^T}(\hat{p}, \tau) = A_i^{\tau}(\hat{p}) \tag{1.12}$$

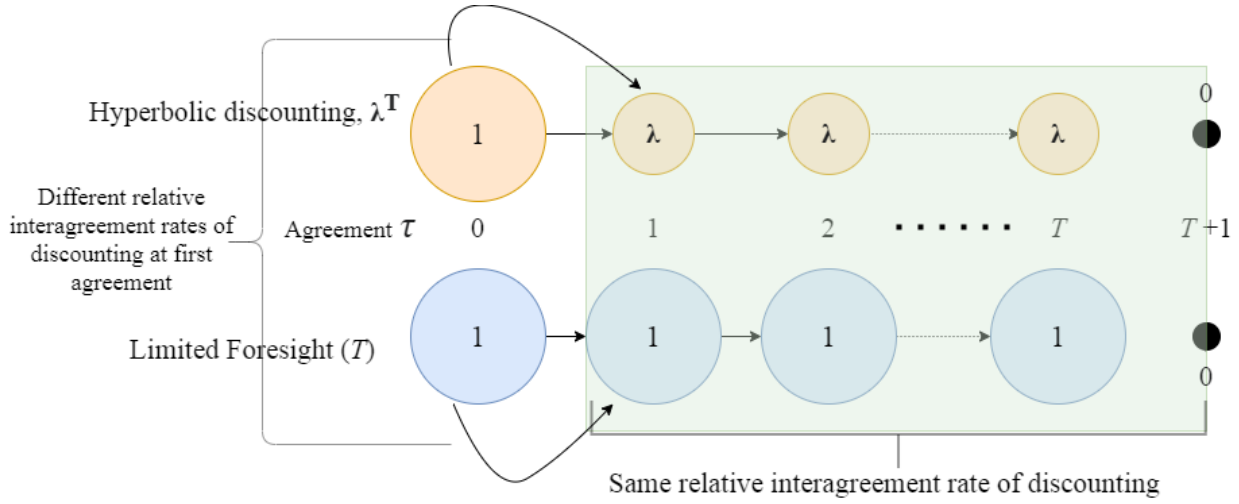


Figure 1.8: A comparison of two different notions of foresight. Note that both hyperbolic discounting in agreements and limited foresight have the same rate of discounting in relevance from one agreement to the other, after the first. Thus, all properties of equilibrium strategies will be the same *after the first* by the principle of backwards induction.

That is, the *estimate* of the strategy player i will use under λ^T- foresight after completing the first agreement are identical to the estimates she made under limited foresight. This is because the *relative* trade-offs across agreements after the first is enacted are the same as with limited foresight (See Figure 1.8). In general,

$$W_R^{\lambda^T}\left(a_L^{\lambda^T}(e,\hat{a})\right) = e\hat{a} + \delta\mathbb{E}\left[W_R^{\lambda^T}\left(a^{\lambda^T}(\tilde{\beta}_1 e,\hat{a})\right)\right]$$
$$W_L^{\lambda^T}\left(a_R^{\lambda^T}(e,\hat{a})\right) = e(1-\hat{a}) + \delta\mathbb{E}\left[W_L^{\lambda^T}\left(a^{\lambda^T}(\tilde{\beta}_1 e,\hat{a})\right)\right]$$

But (1.12) allows us to replace $W_R^{\lambda^T}(a)$ with $\lambda W_R^T(a)$ and thereby also use the identity $W_L^T(a) = S_{T-1} - W_R^T(a)$ from Section 1.4 for any $a \in [0,1]$. We therefore obtain

$$\Delta^{\lambda^T}(e) = \frac{(1-\delta)S_{T-1}}{\mathcal{H}^T} - \frac{e}{\lambda\mathcal{H}^T} \quad \Longrightarrow \quad e_*^{\lambda^T} = \lambda e_T^*.$$

Remark 1. *By Proposition 3,* $\lim_{T\to\infty} e_*^{\lambda^T} = \lambda$. *Thus, if the stochastic process* $\tilde{\beta}$ *has* $F(\lambda) < 1$ *then there will be gridlock on the path of play with probability 1 even in the limiting case where*

$T \to \infty$. *This example demonstrates that even when policy makers consider an arbitrarily large number of agreements, devaluing future agreements can engender gridlock.*

The analysis of Example 1 therefore poses the following question— *what are sufficient conditions on λ for gridlock to be generated under $\lambda-$ foresight?* The main results of Section 3.4, along with this example suggest that if we normalize $\lambda_0 = 1$, then the having *any* $\lambda_t < 1$ is sufficient to induce gridlock. On the other hand, we know from Proposition 3 that this is certainly a necessary condition. This avenue of exploration allows us to greatly generalize the results of the workhorse model.

Differently Foresighted Politicians

Another important question that requires further exploration is how differences in foresight amongst parties can alter negotiations. In particular, while it is well known that the more "patient" (higher δ) party in bargaining is able to leverage this to its benefit in negotiations because its outside option from "holding out" is more attractive— e.g., Rubinstein (1982), the effect for the more foresighted party may well be the opposite. This is because, as Proposition 2 reveals, foresight causes agents to be *more lax* in reaching agreements, since they recognize that there are both future opportunities to change the partisan lean of policy, and because they value downstream efficiency gains more than their counterpart. Hence, one would expect the less foresighted party to be more stubborn in negotiations and extract concessions in the present negotiations from the more forward-looking party.

1.5.2 Concluding Remarks

This paper constructs a dynamic bargaining model based on the assumptions that agreements become less efficient over time and that agents do not incorporate the full stream of future policy consequences into decision-making. The model posits a new source of gridlock, limited foresight on the part of players. In the model, inefficiency decreases as the horizon of foresight increases, and vanishes in the limit as foresight expands to include the full stream of future agreements brokered by

a polity across time. Crucially, I find that the foresight horizon has the opposite effect on gridlock as the intertemporal discount factor, δ.

In our central application, legislative bargaining in Section 1.4, power is also endogenous so as to model electoral constraints faced by politicians in a legislative setting. Within this setting, I show that policy proposals exhibit an *attenuated* ideological persistence, in expectation, over time. The attenuation allows us to characterize the long-run distribution of the partisan lean of policies in equilibrium play as invariant and symmetric about the expected ideology of the median voter, $\frac{1}{2}$. I also study the distance between party's equilibrium proposals and show that this gap increases as foresight *increases*; this is explained by parties' willingness to forgo ideological concerns in the present for efficiency gains, allowing the opposition to exhibit greater permissiveness in the policies she allows to pass.

Bibliography

Acharya, A. and Ortner, J. (2020). Paths to the frontier. *AEJ: Microeconomics*, (forthcoming).

Ali, S. N. M. (2006). Waiting to settle: Multilateral bargaining with subjective biases. *Journal of Economic Theory*, 130(1):109 – 137.

Anesi, V. and Duggan, J. (2018). Existence and indeterminacy of markovian equilibria in dynamic bargaining games. *Theoretical Economics*, 13(2):505 – 525.

Austen-Smith, D., Dziuda, W., Harstad, B., and Loeper, A. (2019). Gridlock and inefficient policy instruments. *Theoretical Economics*, 14(4):1483–1534.

Bai, J. H. and Lagunoff, R. (2011). On the Faustian Dynamics of Policy and Political Power. *The Review of Economic Studies*, 78(1):17–48.

Banks, J. S. and Duggan, J. (2008). A dynamic model of democratic elections in multidimensional policy spaces. *Quarterly Journal of Political Science*, 3(3):269–299.

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

Binder, S. A. (1999). The dynamics of legislative gridlock, 1947-96. *The American Political Science Review*, 93(3):519–533.

Callander, S. and Martin, G. (2020). Policy decay and political competition. *mimeo*.

Callander, S. and Martin, G. J. (2017). Dynamic policymaking with decay. *American Journal of Political Science*, 61(1):50–67.

Duggan, J. (2000). Repeated elections with asymmetric information. *Economics & Politics*, 12(2):109–135.

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

Dziuda, W. and Loeper, A. (2016). Dynamic collective choice with endogenous status quo. *Journal of Political Economy*, 124(4):1148–1186.

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

Gregory, R. (1989). Political rationality or 'incrementalism'?: Charles e. lindblom's enduring contribution to public policy making theory. *Policy and Politics*, 17(2):139–153.

Hayes, M. (2017). Incrementalism and public policy. *Oxford Research Encyclopedia*, Politics(Policy, Administration, and Bureaucracy):1–24.

Jehiel, P. (1995). Limited horizon forecast in repeated alternate games. *Journal of Economic Theory*, 67(2):497 – 519.

Jehiel, P. (1998). Repeated games and limited forecasting. *European Economic Review*, 42(3):543 – 551.

Jehiel, P. (1999). Predicting by analogy and limited foresight in games. *Semantic Scholar*, pages 1–39.

Kalandrakis, A. (2004). A three-player dynamic majoritarian bargaining game. *Journal of Economic Theory*, 116(2):294 – 14.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2):443–477.

Lee, F. E. (2016). *Insecure Majorities: Congress and the Perpetual Campaign*. University of Chicago Press.

Lindblom, C. E. (1959). The science of "muddling through". *Public Administration Review*, 19(2):79–88.

Mansbridge, J. and Martin, C. J. (2013). *Negotiating Agreement in Politics: Report of the Task Force on Negotiating Agreement in Politics*. American Political Science Association. `https://scholar.harvard.edu/files/dtingley/files/negotiating_agreement_in_politics.pdf`(visited 2020-02-08).

Ortner, J. (2013). Optimism, delay and (in)efficiency in a stochastic model of bargaining. *Games and Economic Behavior*, 77(1):352 – 366.

Ortner, J. (2017). A theory of political gridlock. *Theoretical Economics*, 12(2):555 – 586.

Poole, K. T. (2005). *Spatial Models of Parliamentary Voting*. Cambridge University Press.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–109.

Shannon, C. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423.

Wichardt, P. (2010). Modelling equilibrium play as governed by analogy and limited foresight. *Games and Economic Behavior*, 70(2):472–487.

Wittman, D. (1983). Candidate motivation: A synthesis of alternative theories. *American Political Science Review*, 77(1):142–157.

Yildiz, M. (2003). Bargaining without a common prior: An immediate agreement theorem. *Econometrica*, 71(3):793–811.

Yildiz, M. (2011). Bargaining with optimism. *Annual Review of Economics*, 3(1):451–478.

1.A Omitted Proofs for Section 3

Proof. (of Lemma 1)
If $p_I^\tau(\hat{p}) \notin A^\tau(\hat{p})$, then either

(1) $p_I^\tau(\hat{p}) \notin A_{-I}^\tau(\hat{p})$, in which case, the proposal is rejected by definition of A_{-I}.

(2) $p_I^\tau(\hat{p}) \notin A_I^\tau(\hat{p})$. Either the proposal is rejected, as in (1), or, if accepted, yields an inferior continuation value for I than maintaining the status-quo and moving to the next period. The latter follows from the fact that $A_I^\tau(\hat{p})$ maximizes utility in equilibrium and incumbency is not history-dependent.

On the other hand, if $\exists\ (1, a) \in A^\tau(\hat{p})$, then if proposed, it will be accepted and implemented. By (2), the proposer prefers to implement such an a rather than maintain the status-quo; moreover, any proposal $p_I \notin A_{-I}^\tau(\hat{p})$ will be rejected. Hence, in equilibrium, it must be that some $(1, a) \in A^\tau(\hat{p})$ comes into place. □

Proposition 1. For any $T \in \mathbb{N}$ and state $((e, \hat{a}), T)$, every equilibrium of the game is in cut-off strategies with unique threshold $e_T^* = (1 - \delta)S_{T-1}$ for agreement to occur.

Proof. We induct on the foresight horizon T. Let $V_i^T(e, \hat{a}|\sigma)$ be the continuation value of agent i after new status quo $(e, \hat{a})$ has materialized and prior to the selection of the incumbent, given that agents play profile σ and have foresight horizon T. Let also $\mathbb{E}[\tilde{\eta}^\tau]$ be the expected surplus remaining in the game when there are τ agreements remaining on the agents' foresight horizon after an agreement is reached in the current period. We first consider $T = 1$. With status quo $(e, \hat{a})$, the condition for both agents to agree to a given proposal $(1, a^1)$ is

$$u_{x_i}(1, a^1) \cdot \mathbb{E}[\tilde{\eta}^0] \geq e\hat{a} + \delta\mathbb{E}[V_i^1(\tilde{\beta}_1 e, \hat{a}|\sigma)] \tag{1.13}$$

for each $i \in \{L, R\}$, given that future play is described by profile σ. Noting the form of u_{x_i}, summing (1.13) across both agents and re-arranging gives us

$$e + \delta\mathbb{E}[S^1(\tilde{\beta}_1 e, \hat{a}|\sigma)] \leq \mathbb{E}[\tilde{\eta}^0], \tag{1.14}$$

where $S^1(e', \hat{a}|\sigma) := \sum_{i\in\{1,2\}} V_i^1(e', \hat{a}|\sigma)$ is the expected total surplus to both bargaining agents when they play profile σ, with status quo e'. By Lemma 1, σ must dictate (dis)agreement at $(e, \hat{a})$ regardless of the identity of the incumbent. Moreover, σ determines the expected value $S^1(e', \hat{a}|\sigma)$ of projected surplus generated throughout negotiations. Suppose that in the case of current-period disagreement, under σ, agreement is reached $\tilde{t}_1 \geq 1$ (possibly stochastic) periods into the future. The corresponding future surplus is given by

$$\mathbb{E}\left[S^1(\tilde{\beta}_1 e, \hat{a}|\sigma)\right] = \mathbb{E}\left[\left(\sum_{t=1}^{\tilde{t}_1-1} \delta^{t-1} \prod_{i=1}^{t} \tilde{\beta}_i\right) e + \delta^{\tilde{t}_1-1} \frac{1}{1 - \delta\overline{\beta}}\right],$$

so that

$$e + \delta\mathbb{E}\left[S^1(\tilde{\beta}_1 e, \hat{a}|\sigma)\right] = e \cdot \mathbb{E}\left[\sum_{t=0}^{\tilde{t}_1-1} \left(\delta^t \prod_{i=1}^{t} \tilde{\beta}_i\right) \middle| \tilde{t}_1, \sigma\right] + \frac{\mathbb{E}[\delta^{\tilde{t}_1}]}{1 - \delta\overline{\beta}}.$$

Substituting back into (1.14) and simplifying, using the fact that $\mathbb{E}[\tilde{\eta}^1] = \frac{1}{1-\delta\overline{\beta}}$, gives us the condition

$$e \leq \frac{1 - \mathbb{E}[\delta^{\tilde{t}_1}]}{(1-\delta\overline{\beta})\mathbb{E}\left[\sum_{t=0}^{\tilde{t}_1-1}\left(\delta^t \prod_{i=1}^{t}\tilde{\beta}_i\right)\middle|\tilde{t}_1, \sigma\right]}$$

for agreement to occur. Noting that for any profile σ describing future gameplay, the above upper bound is minimal for $\tilde{t}_1 = 1$, it follows that agreement must occur in the current period if $e \leq e^* := \frac{1-\delta}{1-\delta\overline{\beta}}$.

On the other hand, consider $e > e^*$. Under any σ, if agents disagree at $(e, \hat{a})$, there is from above a path of future play in which agreement occurs after $t \geq 1$ periods. In this case,

$$\begin{aligned} e + \delta\mathbb{E}\left[S^1(\tilde{\beta}_1 e, \hat{a}|\sigma, t)\right] &= e + \mathbb{E}\left[\sum_{j=1}^{t-1}\delta^j\left(\prod_{i=1}^{j}\tilde{\beta}_i\right)e\,\middle|\sigma, t\right] + \frac{\delta^t}{1-\delta\overline{\beta}} \\ &= \sum_{j=0}^{t-1}\delta^j\mathbb{E}\left[e\cdot\prod_{i=1}^{j}\tilde{\beta}_i\middle|\sigma, t\right] + \frac{\delta^t}{1-\delta\overline{\beta}} \qquad (1.15) \\ &> \frac{1-\delta}{1-\delta\overline{\beta}}\left(\frac{1-\delta^t}{1-\delta}\right) + \frac{\delta^t}{1-\delta\overline{\beta}} = \frac{1}{1-\delta\overline{\beta}}. \end{aligned}$$

Here, the final inequality results from noting that each element in the sum of (1.15) must have $\mathbb{E}\left[e\cdot\prod_{i=1}^{j}\tilde{\beta}_i\middle|\sigma, t\right] > e^* = \frac{1-\delta}{1-\delta\overline{\beta}}$, since at any future period j ($j < t$) for which $e\cdot\prod_{i=1}^{j}\tilde{\beta}_i \leq e^*$, we have from above that there must be agreement in period j, which would contradict the assumption of t periods of disagreement. The final inequality shows that for any $e > e^*$, condition (1.14) for agreement to occur is violated. So, gridlock occurs at $e > e^*$ in equilibrium.

Hence, defining $e_1^* := \frac{1-\delta}{1-\delta\overline{\beta}} = (1-\delta)S_0$, agreement occurs in any equilibrium if and only if $e \leq e_1^*$ for any status quo partisan lean $\hat{a}$. This establishes that for $T = 1$, each equilibrium is in cut-off strategies, with a unique cut-off in efficiency for agreement to occur.

We now suppose the statement is true for $\tau = T - 1$ and consider the behavior of agents with foresight horizon T. By inductive hypothesis and definition of S_{T-1}, starting at the period in which the next agreement is passed, the remaining expected surplus in the game remains constant at S_{T-1} for any equilibrium strategy profile σ. The proof of the proposition for any $T > 1$ is thus identical to the base case, substituting S_{T-1} for $S_0 = \frac{1}{1-\delta\overline{\beta}}$, $\tilde{t}_T$ for $\tilde{t}_1$, $S^T(\cdot)$ for $S^1(\cdot)$, and $V_i^T(\cdot)$ for $V_i^1(\cdot)$. □

Proposition 2. *In any equilibrium of the game,*

(a) for $T' < T$ and any set of other fixed parameters, $e_{T'}^* < e_T^*$

(b) for any $T \in \mathbb{N}$, e_T^* is strictly decreasing in δ

Proof. (*of Proposition 2*) (a) From Proposition 1, for every T,

$$e_T^* = (1-\delta)S_{T-1} = (1-\delta)\left(\sum_{k=1}^{T}\mathbb{E}\left[\delta^{\sum_{i=1}^{k-1}\tilde{t}_{T-i}}\tilde{\eta}^{T-k}\right]\right) = (1-\delta)\left(\mathbb{E}\left[\tilde{\eta}^{T-1}\right] + \mathbb{E}\left[\delta^{\tilde{t}_{T-1}}S_{T-2}\right]\right)$$

$$
\begin{aligned}
&> (1-\delta)\mathbb{E}\left[\sum_{i=0}^{\tilde{t}_{T-1}} \delta^i \cdot e^*_{T-1} + \delta^{\tilde{t}_{T-1}} S_{T-2}\right] \\
&= (1-\delta) S_{T-2}\mathbb{E}\left[\sum_{i=0}^{\tilde{t}_{T-1}} \delta^i \cdot (1-\delta) + \delta^{\tilde{t}_{T-1}}\right] \\
&= (1-\delta) S_{T-2} = e^*_{T-1},
\end{aligned}
$$

where the final equality results from the fact that $\sum_{i=0}^{\tilde{t}_{T-1}} \delta^i \cdot (1-\delta) + \delta^{\tilde{t}_{T-1}} = 1$ for any integral realization of $\tilde{t}_{T-1}$.

(b) For convenience, let $\tilde{\beta}_0 = 1$. We begin by defining

$$
N(x) := \mathbb{E}\left[\sum_{t\in\mathbb{N}:\prod_{i=0}^{t}\tilde{\beta}_i > x} \delta^t \left(\prod_{i=1}^{t}\tilde{\beta}_i\right)\right], \tag{1.16}
$$

$$
d(x) := \mathbb{E}\left[\delta^{\min\{t\in\mathbb{N}:\prod_{i=1}^{t}\tilde{\beta}_i \le x\}}\right]. \tag{1.17}
$$

$N(x)$ is expected efficiency life of an agreement renegotiated only once $e \le x$ and $d(x)$ is the expected δ-power of the time of renegotiation. These will be helpful objects in our analysis. We consider first e^*_1. It is clear that

$$
\frac{\partial e^*_1}{\partial \delta} = -\frac{1-\overline{\beta}}{(1-\delta\overline{\beta})^2} < 0.
$$

Next, note that for each $T \ge 2$,

$$
\frac{\partial e^*_T}{\partial \delta} = \frac{\partial\{(1-\delta)S_{T-1}\}}{\partial \delta} = -S_{T-1} + (1-\delta)\frac{\partial}{\partial\delta}S_{T-1}. \tag{1.18}
$$

The following auxiliary claim is thus useful in completing the proof—

Lemma 3. *For each $T \in \mathbb{N}$,*

$$
\frac{\partial}{\partial\delta}S_T = \frac{\prod_{\tau=1}^{T} d(e^*_\tau)\cdot\overline{\beta}}{(1-\delta\overline{\beta})^2}. \tag{1.19}
$$

Proof of Lemma 3. We proceed by induction on T. First, note that using definitions (1.16) and (1.17), we can write

$$
S_{T-1} = N(e^*_{T-1}) + d(e^*_{T-1})S_{T-2}.
$$

Thus, for the base case $T = 1$,

$$
\begin{aligned}
\frac{\partial}{\partial\delta}S_1 &= \frac{\partial}{\partial\delta}\left(N(e^*_1) + d(e^*_1)\cdot\frac{1}{1-\delta\overline{\beta}}\right) \\
&= N'(e^*_1)\cdot\frac{\partial}{\partial\delta}e^*_1(\delta) + d'(e^*_1)\cdot\frac{\partial}{\partial\delta}e^*_1(\delta)\cdot\frac{1}{1-\delta\overline{\beta}} + d(e^*_1)\cdot\frac{\partial}{\partial\delta}\frac{1}{1-\delta\overline{\beta}}
\end{aligned}
$$

$$= -\frac{1}{(1-\delta\overline{\beta})^2}\left[(1-\overline{\beta})\cdot\left(N'(e_1^*)+\frac{d'(e_1^*)}{1-\delta\overline{\beta}}\right)-d(e_1^*)\overline{\beta}\right] \tag{1.20}$$

We evaluate by rewriting $N(e_1^*)$ and $d(e_1^*)$ in convenient forms for differentiation. For convenience, we let G_i be the cumulative distribution of the stochastic product $\prod_{j=1}^{i}\tilde{\beta}_j$ and g_i its pdf. For any x,

$$\begin{aligned}N(x) &= 1+\delta\cdot\mathbb{P}[\tilde{\beta}_1>x]\cdot\mathbb{E}[\tilde{\beta}_1|\tilde{\beta}_1>x]+\delta^2\cdot\mathbb{P}[\tilde{\beta}_1>x]\cdot\mathbb{E}[\tilde{\beta}_1|\tilde{\beta}_1>x]+\dots\\ &= 1+\delta\cdot\int_x^1 t\cdot g_1(t)\,dt+\delta^2\cdot\int_x^1 t\cdot g_2(t)\,dt+\dots.\end{aligned}$$

Using the Leibniz rule,

$$N'(x) = -x\cdot\sum_{i=1}^{\infty}\delta^i g_i(x). \tag{1.21}$$

Similarly,
$$\begin{aligned}&d(x) = \delta G_1(x)+\delta^2(G_2(x)-G_1(x))+\delta^3(G_3(x)-G_2(x))+\dots\\ \implies\ &d'(x) = \delta g_1(x)+\delta^2(g_2(x)-g_1(x))+\delta^3(g_3(x)-g_2(x))+\dots.\end{aligned}$$

By directly applying the above expressions, we then obtain

$$N'(e_1^*)+\frac{d'(e_1^*)}{1-\delta\overline{\beta}} = \sum_{i=1}^{\infty}\left[g_i(e_1^*)\left(-\delta^i e_1^*+\frac{\delta^i}{1-\delta\overline{\beta}}-\frac{\delta^{i+1}}{1-\delta\overline{\beta}}\right)\right] = 0,$$

from the fact that $e_1^* = \frac{1-\delta}{1-\delta\overline{\beta}}$. Therefore, substituting back into (1.20),

$$\frac{\partial}{\partial\delta}S_1 = \frac{d(e_1^*)\overline{\beta}}{(1-\delta\overline{\beta})^2}.$$

This shows the base case. We suppose now the claim holds true for any $\tau = 1,\dots,T-1$. We show it must hold for T.

$$\begin{aligned}S_T &= N(e_T^*)+d(e_T^*)S_{T-1}.\\ \implies \frac{\partial}{\partial\delta}S_T &= N'(e_T^*)\cdot\frac{\partial e_T^*}{\partial\delta}+d'(e_T^*)\cdot\frac{\partial e_T^*}{\partial\delta}S_{T-1}+d(e_T^*)\cdot\frac{\partial}{\partial\delta}S_{T-1}\\ &= \frac{\partial e_T^*}{\partial\delta}\left(N'(e_T^*)+d'(e_T^*)S_{T-1}\right)+d(e_T^*)\cdot\frac{\partial}{\partial\delta}S_{T-1}\\ &= \frac{\partial e_T^*}{\partial\delta}\cdot\sum_{i=1}^{\infty}\left[g_i(e_T^*)\left(-\delta^i e_T^*+(\delta^i-\delta^{i+1})S_{T-1}\right)\right]+d(e_T^*)\cdot\frac{\partial}{\partial\delta}S_{T-1},\end{aligned}$$

where in the final equality we reapply the expressions for $N'(x)$ and $d'(x)$ derived above. Since $e_T^* = (1-\delta)S_{T-1}$, each summand $g_i(e_T^*)\left(-\delta^i e_T^*+(\delta^i-\delta^{i+1})S_{T-1}\right) = 0$. Hence,

$$\frac{\partial}{\partial\delta}S_T = d(e_T^*)\cdot\frac{\partial}{\partial\delta}S_{T-1}$$

$$= d(e_T^*)\frac{\prod_{\tau=1}^{T-1} d(e_*^\tau)\overline{\beta}}{(1-\delta\overline{\beta})^2}, \text{ using the inductive hypothesis for } \frac{\partial}{\partial\delta}S_{T-1}$$
$$= \frac{\prod_{\tau=1}^{T} d(e_*^\tau)\overline{\beta}}{(1-\delta\overline{\beta})^2}, \text{ as desired.}$$

□

Using Lemma 3, we can now complete the proof of Proposition 2(b) for any $T \geq 2$.

$$\frac{\partial}{\partial\delta}e_*^T = -S_{T-1} + (1-\delta)\frac{\partial}{\partial\delta}S_{T-1} = -S_{T-1} + \frac{(1-\delta)\prod_{\tau=1}^{T} d(e_*^\tau)\overline{\beta}}{(1-\delta\overline{\beta})^2} < 0,$$

where the concluding inequality results from the fact that $\frac{1-\delta}{1-\delta\overline{\beta}} < 1$ and $S_{T-1} > S_0 := \frac{1}{1-\delta\overline{\beta}}$ since from part (a), the S_T are increasing in T. This completes the proof. □

Proposition 3. *Let* $\lim_{T\to\infty} e_T^* := e_\infty^*$. *Then,* $e_\infty^* = 1$*; there is no delay on equilibrium path for any stochastic process* $\tilde{\beta}$.

Proof. Recall the definitions (1.16), (1.17) from above. $N(x)$ is the expected efficiency life of an agreement renegotiated only once $e \leq x$ and $d(x)$ is the expected δ-power of when this agreement is renegotiated. By definition,

$$\begin{aligned} e_T^* &= (1-\delta)S_{T-1} \\ &= (1-\delta)\left\{N(e_*^{T-1}) + d(e_*^{T-1})N(e_*^{T-2}) + d(e_*^{T-2})d(e_*^{T-1})N(e_*^{T-3}) + \ldots\right\}, \end{aligned} \tag{1.22}$$

in terms of the functions $N(\cdot)$ and $d(\cdot)$. From the fact that $\tilde{\beta}$ has a non-atomic distribution, both are continuous in x and therefore also uniformly continuous on the domain $[0,1]$. Thus, for any $\varepsilon > 0$, $\exists\varphi_1, \varphi_2$ such that

$$|x-y| < \varphi_1 \implies |N(x) - N(y)| < \frac{\varepsilon}{2}, \quad |x-y| < \varphi_2 \implies |d(x) - d(y)| < \frac{\varepsilon}{2}.$$

Letting $\varphi := \max\{\varphi_1, \varphi_2\}$ gives us

$$|x-y| < \varphi \implies |N(x) - N(y)| < \varepsilon, \quad |d(x) - d(y)| < \varepsilon. \tag{1.23}$$

Consider arbitrary $\varepsilon \in (0, 2\delta)$ and the corresponding φ. From Proposition 2(a), the sequence $\{e_T^*\}$ is monotonically increasing and bounded from above (by 1) and so (Cauchy) converges. Hence, there exists T_φ such that $\forall\ m, n > T_\varphi$, $|e_*^m - e_*^n| < \varphi$. Consider the expression (1.22) for $2T_\varphi$:

$$e_*^{2T_\varphi} = (1-\delta)\left\{N\left(e_*^{2T_\varphi-1}\right) + d\left(e_*^{2T_\varphi-1}\right)N\left(e_*^{2T_\varphi-2}\right) + d\left(e_*^{2T_\varphi-1}\right)\cdot d\left(e_*^{2T_\varphi-2}\right)N\left(e_*^{2T_\varphi-3}\right) + \ldots\right\}.$$

Hence, by using (1.23),

$$\left|e_*^{2T_\varphi} - (1-\delta)N\left(e*^{2T_\varphi}\right)\sum_{i=0}^{2T_\varphi} d\left(e_*^{2T_\varphi}\right)\right| \leq (1-\delta)\left(\sum_{i=1}^{T_\varphi}\left(\frac{\varepsilon}{2}\right)^i + \frac{\delta^{T_\varphi}}{1-\delta}\right)$$

$$\leq (1-\delta)\cdot\frac{\varepsilon}{2-\varepsilon}+\delta^{T_\varphi}$$
$$\leq \frac{\varepsilon}{2}+\delta^{T_\varphi},\ \text{since}\ \frac{\varepsilon}{2}<\delta.$$

Now, since the statement holds for all T_φ sufficiently large, we can simply choose $T_\varepsilon > T_\varphi$ in the above analysis so that $\delta^{T_\varepsilon} < \frac{\varepsilon}{2}$. Since $\delta < 1$, such a T_ε exists. The above shows that for any $\varepsilon > 0$, there exists $T_\varepsilon > 0$ such that for all $T > T_\varepsilon$,

$$\left| e_*^{2T_\varepsilon} - (1-\delta)N\left(e*^{2T_\varepsilon}\right)\sum_{i=0}^{2T_\varepsilon} d\left(e_*^{2T_\varepsilon}\right)\right| \leq \varepsilon.$$

Hence, in the limit as $T \to \infty$,

$$\lim_{T\to\infty} e_T^* = e_\infty^* = (1-\delta)N\left(e_\infty^*\right)\frac{1}{1-d\left(e_\infty^*\right)}$$
$$= (1-\delta)\mathbb{E}\left[\sum_{t\in\mathbb{N}:\prod_{i=1}^t \tilde{\beta}_i > e_\infty^*} \delta^t\left(\prod_{i=1}^t \tilde{\beta}_i\right)\right]\frac{1}{1-\mathbb{E}\left[\delta^{\min\{t\in\mathbb{N}:\prod_{i=1}^t\tilde{\beta}_i\leq e_\infty^*\}}\right]}, \quad (1.24)$$

substituting in the definitions (1.16) and (1.17). Let β^* (possibly equal to 1) be the supremum of the support of the random variable $\tilde{\beta}$. First, note that $e_\infty^* = 1$ solves the fixed point equation (1.24) since the right-hand side is 1 for any $e_\infty^* \in [\beta^*, 1]$. If $\beta^* < 1$, this also rules out any $e_\infty^* \in [\beta^*, 1)$ as a solution. Furthermore, for any sequence $(x_i)_{i=0}^\infty$, with $x_i \in (e_\infty^*, 1)$,

$$(1-\delta)\frac{1+\delta x_1+\delta^2 x_2+\ldots+\delta^t x_t}{1-\delta^{t+1}} = \frac{1+\delta x_1+\delta^2 x_2+\ldots+\delta^t x_t}{1+\delta+\ldots+\delta^t} > \max_t x_t > e_\infty^*,$$

for any $t \in \mathbb{N}$. Hence, for any realization of $\{\tilde{\beta}_i\}$ such that $\min\{t\in\mathbb{N}:\prod_{i=1}^t\tilde{\beta}_i \leq e_\infty^*\} > 1$, it follows that

$$(1-\delta)\left[\sum_{t\in\mathbb{N}:\prod_{i=1}^t\tilde{\beta}_i > e_\infty^*}\delta^t\left(\prod_{i=1}^t\tilde{\beta}_i\right)\right]\frac{1}{1-\delta^{\min\{t\in\mathbb{N}:\prod_{i=1}^t\tilde{\beta}_i\leq e_\infty^*\}}} > e_\infty^*.$$

Otherwise, it is equal to 1. Noting that for any $e_\infty^* < \beta^*$ there will be delay with positive probability before agreement, it follows that the right-hand side of (1.24) is greater than e_∞^* for any such e_∞^* and thus cannot solve (1.24). Hence, $e_\infty^* = 1$ is the unique solution to (1.24), completing the proof.

□

1.B Omitted Proofs for Section 1.4

Note: for ease of notation in the proofs in this Appendix, we drop σ from the terms $p^T(\hat{p}, I_T, \sigma)$ and $a^T(\hat{p}, I_t, \sigma)$ for outcomes that materialize in equilibrium given a particular status quo and incumbent.

Proof of Lemma 2. By definition, at any status quo $(e, \hat{a})$ at which agreement occurs, $\mu^T(e, \hat{a})$ solves

$$-\mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \frac{e_{\tilde{q}_t^T(R)} \cdot (\mu^T(e,\hat{a}) - a_{\tilde{q}_t^T(R)})^2}{|a_{\tilde{q}_t^T(R)} - a_{\tilde{q}_t^T(L)}|}\right] = -\mathbb{E}\left[\delta^t \sum_{t=0}^{\infty} \frac{e_{\tilde{q}_t^T(L)} \cdot (\mu^T(e,\hat{a}) - a_{\tilde{q}_t^T(L)})^2}{|a_{\tilde{q}_t^T(R)} - a_{\tilde{q}_t^T(L)}|}\right],$$

which implies

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t \left(\frac{e_{\tilde{q}_t^T(L)} \cdot (\mu^T(e,\hat{a}) - a_{\tilde{q}_t^T(L)})^2}{|a_{\tilde{q}_t^T(R)} - a_{\tilde{q}_t^T(L)}|} - \frac{e_{\tilde{q}_t^T(R)} \cdot (\mu^T(e,\hat{a}) - a_{\tilde{q}_t^T(R)})^2}{|a_{\tilde{q}_t^T(R)} - a_{\tilde{q}_t^T(L)}|}\right)\right] = 0.$$

From Lemma 1 and Assumption 2, $e_{\tilde{q}_t^T(R)} = e_{\tilde{q}_t^T(L)}$, so that we can simply write e_t. Through factorization, the above reduces to

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \delta^t e_t \frac{(2\mu^T(e,\hat{a}) - a_{\tilde{q}_t^T(R)} - a_{\tilde{q}_t^T(L)}) \cdot (a_{\tilde{q}_t^T(R)} - a_{\tilde{q}_t^T(L)})}{|a_{\tilde{q}_t^T(R)} - a_{\tilde{q}_t^T(L)}|}\right] = 0.$$

And since, by hypothesis, $a_{\tilde{q}_t^T(R)} > a_{\tilde{q}_t^T(L)}$ for any arbitrary realization of future states, we obtain

$$\mu^T(e,\hat{a}) = \frac{\sum_{t=0}^{\infty} \delta^t e_t \cdot (a_{\tilde{q}_t^T(R)} + a_{\tilde{q}_t^T(L)})}{2\sum_{t=0}^{\infty} \delta^t e_t} = \frac{W_R^T(a_R^T(e,\hat{a})) + W_R^T(a_L^T(e,\hat{a}))}{2S_{T-1}}.$$

□

1.B.1 Proofs for Section 1.4.2

Proposition 4. For any $T \in \mathbb{N}$, there is an almost-everywhere outcome-unique equilibrium of the game that implements unique proposals $p_i^T(\hat{p})$ with $e_i^T(\hat{p}) = 1$ and $a_i^T(e,\hat{a})$ linear and increasing in $\hat{a}$ for $i \in \{R, L\}$ for any $\hat{p} = (e, \hat{a})$ at which agreement occurs on the equilibrium path. We write for each i,

$$a_i^T(e,\hat{a}) = C^T(e)\hat{a} + B_i^T(e),$$

where $C^T \in (0,1)$ for $e > 0$ and $C^T(0) = 0$.

Proof. We prove the result by induction on the foresight horizon T through two lemmas— a preliminary result followed by the inductive step. The base case of $T = 1$ is in the main text. We will first show that if the main proposition holds for all $\tau \leq T-1$, that linearity in the partisan lean of policy, $\hat{a}$, also holds for several key equilibrium objects.

Let $\mathcal{H}^1 := \frac{1}{1-\delta\bar{\beta}}$ and $\mathcal{J}^1 := 0$. Supposing Proposition 4 holds for $\tau \leq T-1$, define recursively for all such τ—

$$K^\tau(e) = C^\tau(e) \cdot \left(1 - \frac{\mathcal{H}^\tau}{S_{\tau-1}}\Delta^\tau(e)\right), \text{ for any } e \leq e_*^\tau$$

$$M^\tau(e) = B_L^\tau(e) \cdot \left(1 - \frac{\mathcal{H}^\tau}{S_{\tau-1}}\Delta^\tau(e)\right) + \left(1 - \frac{\mathcal{H}^\tau\Delta^\tau(e) + 2\mathcal{J}^\tau}{2S_{\tau-1}}\right)\Delta^\tau(e), \text{ for any } e \leq e_*^\tau$$

$$\mathcal{H}^{\tau+1} = \mathbb{E}[\tilde{\eta}^\tau] + \mathbb{E}[\delta^{\tilde{t}_\tau} \cdot \mathcal{H}^\tau \cdot K^\tau(\tilde{e}^\tau)]$$
$$\mathcal{J}^{\tau+1} = \mathbb{E}\left[\delta^{\tilde{t}_\tau}\left(\mathcal{H}^\tau \cdot M^\tau(\tilde{e}^\tau) + \mathcal{J}^\tau\right)\right]$$

Lemma 4. *Suppose the statement of Proposition 4 holds for every $\tau \leq T-1$ Then, for each $\tau \leq T-1$,*

(i) If I^τ is the incumbent when p^τ is implemented, then for any $e \leq e^\tau_*$,

$$\mathbb{E}_{I^\tau(e)}\left[a^\tau(e,\hat{a})\right] = K^\tau(e)\cdot\hat{a} + M^\tau(e)$$

(ii) Moreover,

$$W_R^{\tau+1}(a) = \mathcal{H}^{\tau+1}a + \mathcal{J}^{\tau+1}.$$

Proof. We proceed by induction on τ. By hypothesis, since $a_i^\tau(e,\hat{a})$ are linear in $\hat{a}$ with the same coefficient $C^\tau(e)$, it follows that $a_R^\tau(e,\hat{a}) - a_L^\tau(e,\hat{a}) = \Delta^\tau(e)$, which does not depend on $\hat{a}$. For $\tau = 1$, we have

(i)
$$\begin{aligned}
\mathbb{E}_{I_1(e)}\left[a^1((e,\hat{a}),I_1)\right] &= a_L^1(e,\hat{a}) + \left(1 - \frac{a_L^1(e,\hat{a}) + a_R^1(e,\hat{a})}{2}\right)\Delta^1(e)\\
&= a_L^1(e,\hat{a}) + \left(1 - a_L^1(e,\hat{a}) - \frac{\Delta^1(e)}{2}\right)\Delta^1(e)\\
&= (1-\Delta^1(e))\cdot a_L^1(e,\hat{a}) + \left(1 - \frac{\Delta^1(e)}{2}\right)\Delta^1(e)\\
&= \underbrace{(1-\Delta^1(e))C^1(e)}_{:=K^1(e)}\cdot a + \underbrace{(1-\Delta^1(e))B_L^1(e) + \left(1-\frac{\Delta^1(e)}{2}\right)\Delta^1(e)}_{:=M^1(e),\quad \text{by definition.}}
\end{aligned}$$

(ii)
$$\begin{aligned}
W_R^2(a) &= a\cdot\mathbb{E}[\tilde{\eta}^1] + \mathbb{E}\left[\delta^{\tilde{t}_1}a^1((\tilde{e}^1,\hat{a}),I_1)\right]\cdot\frac{1}{1-\delta\overline{\beta}}\\
&= a\underbrace{\left(\mathbb{E}[\tilde{\eta}^1] + \mathbb{E}[\delta^{\tilde{t}_1}\mathcal{H}^1K^1(\tilde{e}^1)]\right)}_{:=\mathcal{H}^2} + \underbrace{\mathbb{E}\left[\delta^{\tilde{t}_1}\mathcal{H}^1M^1(\tilde{e}^1)\right]}_{:=\mathcal{J}^2,\quad\text{by definition.}}
\end{aligned}$$

Now suppose (i) and (ii) hold for all $\tau < \hat{T} \leq T-1$. First, we show that (i) holds for $\hat{T}$. Note by inductive hypothesis that $W_R^{\hat{T}}$ is increasing. Thus, parties will push their proposals to the extrema of the acceptability set $A^{\hat{T}}(e,\hat{a})$ in equilibrium, so that the hypothesis of Lemma 2 is satisfied. We then have

$$\begin{aligned}
\mathbb{E}_{I^{\hat{T}}(e)}\left[a^{\hat{T}}((e,\hat{a}),I_{\hat{T}})\right] &= a_L^{\hat{T}}(e,\hat{a}) + \left(1 - \frac{W_R^{\hat{T}}(a_L^{\hat{T}}(e,\hat{a})) + W_R^{\hat{T}}(a_R^{\hat{T}}(e,\hat{a}))}{2S_{T-1}}\right)\Delta^{\hat{T}}(e)\\
&= a_L^{\hat{T}}(e,\hat{a}) + \left(1 - \frac{2\mathcal{H}^{\hat{T}}\cdot a_L^{\hat{T}}(e,\hat{a}) + \mathcal{H}^{\hat{T}}\Delta^{\hat{T}}(e) + 2\mathcal{J}^{\hat{T}}}{2S_{\hat{T}-1}}\right)\Delta^{\hat{T}}(e),
\end{aligned}$$

by inductive hypothesis on (ii)

$$= \left(1 - \frac{\mathcal{H}^{\hat{T}}}{S_{\hat{T}-1}}\Delta^{\hat{T}}(e)\right) a_L^{\hat{T}}(e,\hat{a}) + \left(1 - \frac{\mathcal{H}^{\hat{T}}\Delta^{\hat{T}+2\mathcal{J}^{\hat{T}}}(e)}{2S_{\hat{T}-1}}\right)\Delta^{\hat{T}}(e)$$

$$= \underbrace{\left(1 - \frac{\mathcal{H}^{\hat{T}}}{S_{\hat{T}-1}}\Delta^{\hat{T}}(e)\right) C^{\hat{T}}(e)}_{:= K^{\hat{T}}(e)} \cdot \hat{a} +$$

$$\underbrace{\left(1 - \frac{\mathcal{H}^{\hat{T}}}{S_{\hat{T}-1}}\Delta^{\hat{T}}(e)\right) B_L^{\hat{T}}(e) + \left(1 - \frac{\mathcal{H}^{\hat{T}}\Delta^{\hat{T}+2\mathcal{J}^{\hat{T}}}(e)}{2S_{\hat{T}-1}}\right)\Delta^{\hat{T}}(e)}_{:= M^{\hat{T}}(e), \quad \text{by definition.}}$$

And, for (ii), $W_R^{\hat{T}+1}(a) = a \cdot \mathbb{E}\left[\tilde{\eta}^{\hat{T}}\right] + \mathbb{E}\left[\delta^{\tilde{t}_{\hat{T}}} W_R^{\hat{T}}\left(a^{\hat{T}}((\tilde{e}^{\hat{T}}, a), I_{\hat{T}})\right)\right]$

$$= a \cdot \mathbb{E}\left[\tilde{\eta}^{\hat{T}}\right] + \mathbb{E}\left[\delta^{\tilde{t}_{\hat{T}}}\left\{\mathcal{H}^{\hat{T}}\left(K^{\hat{T}}(\tilde{e}^{\hat{T}}) \cdot a + M^{\hat{T}}(\tilde{e}^{\hat{T}})\right) + \mathcal{J}^{\hat{T}}\right\}\right],$$

by composing linear functions

$$= a \cdot \underbrace{\left(\mathbb{E}\left[\tilde{\eta}^{\hat{T}}\right] + \mathbb{E}\left[\delta^{\tilde{t}_{\hat{T}}}\mathcal{H}^{\hat{T}} \cdot K^{\hat{T}}(\tilde{e}^{\hat{T}})\right]\right)}_{= \mathcal{H}^{\hat{T}+1}} + \underbrace{\mathbb{E}\left[\delta^{\tilde{t}_{\hat{T}}}\left(\mathcal{H}^{\hat{T}} \cdot M^{\hat{T}}(\tilde{e}^{\hat{T}}) + \mathcal{J}^{\hat{T}}\right)\right]}_{= \mathcal{J}^{\hat{T}+1}}.$$

This completes the inductive step and thus the proof. □

Inductive Step. We now use Lemma 4 to complete the inductive step. Take any $T \geq 2$. Suppose the statement of Proposition 4 holds for every $\tau \leq T-1$. Then, the statement also holds for T.

Given the inductive hypothesis, there is an outcome-unique equilibrium in cut-off strategies for foresight horizon of length $T-1$. Let σ be the profile played therein. As before, we continue to write $\tilde{\eta}^{\tau}$ to be the efficiency life of the agreement which is in place while τ future agreements remain on the foresight horizon, and let $\tilde{t}_{\tau}$ be the number of periods this agreement is in place. It remains to be shown that the proposal functions with a T foresight horizon are linear and unique in any equilibrium.

The argument is analogous to the proof of the base case in the main text. Consider $e \leq (1-\delta)S_{T-1}$. As in (1.13), a new agreement a is acceptable to both players at $\hat{p} = (e, \hat{a})$ if and only if

$$W_i^T(a) \geq u_{x_i}(\hat{p}) + \delta\mathbb{E}[V_i^T(\tilde{\beta}_1 e, \hat{a}|\sigma)].$$

Let $y_R^T(e,\hat{a})$ and $y_L^T(e,\hat{a})$ be the minimal (maximal) policies accepted by R and L, respectively whenever $e \leq (1-\delta)S_{T-1}$. We have then

$$W_R^T(y_R^T(e,\hat{a})) = e\hat{a} + \delta\mathbb{E}[V_R^T(\tilde{\beta}_1 e, \hat{a}|\sigma)] \qquad W_L(y_L^T(e,\hat{a}) = e(1-\hat{a}) + \delta\mathbb{E}[V_L^T(\tilde{\beta}_1 e, \hat{a}|\sigma)] \tag{1.25}$$

From Proposition 1, agreement occurs in any equilibrium in the ensuing period should it not occur in the current one. Hence, $\sum_{i\in\{R,L\}} V_i^T(\tilde{\beta}_1 e, \hat{a}|\sigma) = S_{T-1}$ for any realization of $\tilde{\beta}_1$. We can

thus express the equation for L in (1.25) as

$$S_{T-1} - W_R^T(y_L^T(e,\hat{a})) = e - e\hat{a} + \delta\left(S_{T-1} - \mathbb{E}\left[\hat{V}_R^T(\tilde{\beta}_1 e, \hat{a})\right]\right)$$

$$W_R^T(y_L^T(e,\hat{a})) = (1-\delta)S_{T-1} - e + W_R^T(y_R^T(e,\hat{a}), \tag{1.26}$$

by substitution from (1.25)

$$\implies W_R^T(y_L^T(e,\hat{a})) - W_R^T(y_R^T(e,\hat{a})) = (1-\delta)S_{T-1} - e \tag{1.27}$$

Finally, by Lemma 4, linearity of $W_R^T(\cdot)$ allows us to rewrite the left-hand side of (1.27) as

$$\mathcal{H}^T \cdot (y_L^T(e,\hat{a}) - y_R^T(e,\hat{a})) = (1-\delta)S_{T-1} - e.$$

By Lemma 4 and indutive hypothesis, W_R^T is increasing so that R (L) will prefer greater (lesser) agreements. That is, $a_R^T(e,\hat{a}) = y_L^T(e,\hat{a})$ and $a_L^T(e,\hat{a}) = y_R^T(e,\hat{a})$. This immediately yields

$$\Delta^T(e) := a_R^T(e,\hat{a}) - a_L^T(e,\hat{a}) = \frac{(1-\delta)S_{T-1} - e}{\mathcal{H}^T}, \tag{1.28}$$

a function only of e. The previously proven structure of cut-off equilibrium with foresight horizon T (Proposition 1), then allows us to derive $a_L^T(e,\hat{a})$ as the solution to $W_R^T(a_L^T(e,\hat{a}) = e\hat{a} + \delta\mathbb{E}\left[V_R^T(\beta_1 e, \hat{a})\right]$. Because W_R^T is increasing by inductive hypothesis from Lemma 4, we can apply the result of Lemma 2 in the calculations below. Using again Lemma 4,

$$W_R^T(a_L^T(e,\hat{a})) = e\hat{a} + \delta\mathbb{E}\left[W_R^T(a^T((\tilde{\beta}_1 e, \hat{a}), I_1))\right], \text{ since we have a cut-off equilibrium. Thus,}$$

$$\begin{aligned}\mathcal{H}^T a_L^T(e,\hat{a}) + \mathcal{J}^T &= e\hat{a} + \delta\mathbb{E}\left[W_R^T(a_L^T(\tilde{\beta}_1 e,\hat{a})) + \left(1 - \frac{W_R^T(a_L^T(\tilde{\beta}_1 e,\hat{a})) + W_R^T(a_R^T(\tilde{\beta}_1 e,\hat{a}))}{2S_{T-1}}\right)\mathcal{H}^T \cdot \Delta^T(\tilde{\beta}_1 e)\right] \\ &= e\hat{a} + \delta\mathbb{E}\left[\mathcal{H}^T \cdot a_L^T(\tilde{\beta}_1 e,\hat{a})\left(1 - \frac{\mathcal{H}^T}{S_{T-1}}\Delta^T(\tilde{\beta}_1 e)\right) + \left(1 - \frac{\mathcal{H}^T\Delta^T(\tilde{\beta}_1 e) + 2\mathcal{J}^T}{2S_{T-1}}\right)\mathcal{H}^T\Delta^T(\tilde{\beta}_1 e)\right] \\ &\quad + \delta\mathcal{J}^T\end{aligned}$$

Rearranging gives us

$$a_L^T(e,\hat{a}) = \frac{e\hat{a}}{\mathcal{H}^T} + \delta\mathbb{E}\left[\left(1 - \frac{\mathcal{H}^T\Delta^T(\tilde{\beta}_1 e) + 2\mathcal{J}^T}{2S_{T-1}}\right)\Delta^T(\tilde{\beta}_1 e)\right] - (1-\delta)\frac{\mathcal{J}^T}{\mathcal{H}^T} + \delta\mathbb{E}\left[\left(1 - \frac{\mathcal{H}^T}{S_{T-1}}\Delta^T(\tilde{\beta}_1 e)\right)a_L^T(\tilde{\beta}_1 e,\hat{a})\right].$$

Note that the above equation has a $a_L^T(\tilde{\beta}_1 e,\hat{a})$ term on the right-hand side. This gives us a recursive formulation of $a_L^T(\tilde{\beta}_1 e,\hat{a})$. Reapplying this recursive formula allows us to obtain

$$\begin{aligned}a_L^T(e,\hat{a}) = &\underbrace{\left(\sum_{i=0}^{\infty}\delta^i \cdot \mathbb{E}\left[\prod_{j=1}^{i}\left(1 - \frac{\mathcal{H}^T}{S_{T-1}}\Delta^T\left(\prod_{k=1}^{j}\tilde{\beta}_k \cdot e\right)\right)\tilde{\beta}_j\right]\right)\frac{e}{\mathcal{H}^T}}_{C^T(e)} \cdot\hat{a} + \\ &\sum_{i=1}^{\infty}\delta^i \cdot \mathbb{E}\left[\prod_{j=1}^{i-1}\left(1 - \frac{\mathcal{H}^T}{S_{T-1}}\Delta^T\left(\prod_{k=1}^{j}\tilde{\beta}_k \cdot e\right)\right)\cdot\right. \\ &\left.\left\{\left(1 - \frac{\mathcal{H}^T\Delta^T\left(\prod_{k=1}^{i}\tilde{\beta}_k \cdot e\right) + 2\mathcal{J}^T}{2S_{T-1}}\right)\Delta^T\left(\prod_{k=1}^{i}\tilde{\beta}_k \cdot e\right) - \frac{(1-\delta)\mathcal{J}^T}{\delta\mathcal{H}^T}\right\}\right].\end{aligned} \tag{1.29}$$

We let the second term in the above expression be $B_L^T(e)$. As before, $a_R^T(e,\hat{a}) = a_L^T(e,\hat{a}) + \Delta^T(e)$. Thus, both parties' equilibrium proposals are uniquely determined by (1.29), linear, and strictly increasing in $\hat{a}$ for $e > 0$, which is readily seen from the fact that $C^T(e) > 0$ for all $e > 0$. Moreoever, the sum in the coefficient of $C^T(e)$ is seen to be less than $\mathcal{H}^T$, by noting the recursive definitions in the prelude to Lemma 4, showing that $C^T(e) < 1$. This completes the proof.

□

Proof. (*of Corollary 3*)

$$\mathbb{P}_t[\text{R wins}] = \mathbb{P}\left[m_t > \frac{W_R^T(a_L^T(e,\hat{a})) + W_R^T(a_R^T(e,\hat{a}))}{2S_{T-1}}\right] = 1 - \frac{W_R^T(a_L^T(e,\hat{a})) + W_R^T(a_R^T(e,\hat{a}))}{2S_{T-1}},$$

where m_t is the position of the t-period median voter. The first equality is a direct application of Lemma 2 and the final equality results from the assumption that m_t is distributed uniformly on $[0,1]$ (Assumption 3). By Lemma 4, $W_R^T(a_I^T(e,\hat{a}))$ is increasing in $\hat{a}$ for $I \in \{L,R\}$, so that $\mathbb{P}_t$[R wins] is decreasing in $\hat{a}$. The result for $\mathbb{E}[a^T(e,\hat{a})]$ is direct from Lemma 4, noting that the definition of $K^T(e)$ means that $K^T \geq 0$ since $C^T \geq 0$ from Proposition 4. In both cases, monotonicity is strict when $e > 0$, again from direct inspection of $K^T(e)$ and $C^T(e)$. □

Proposition 5 Consider $T' < T$. Whenever $e \leq e_{T'}^* < e_T^*$, $\Delta^T(e) > \Delta^{T'}(e)$. Moreover, $\Delta^T(e)$ is strictly decreasing in δ.

Proof. We prove the first part of the proposition first. To show this, note that it suffices to show $\Delta^{T+1}(e) > \Delta^T(e)$ for $e \leq e_*^T < e_{T+1}^*$. By (1.28) (See, proof of Lemma 1.B.1),

$$\Delta^\tau(e) = \frac{(1-\delta)S_{\tau-1} - e}{\mathcal{H}^\tau}.$$

Define

$$\varphi(e) := \Delta^{T+1}(e) - \Delta^T(e) = (1-\delta)\left(\frac{S_T}{\mathcal{H}^{T+1}} - \frac{S_{T-1}}{\mathcal{H}^T}\right) + \left(\frac{1}{\mathcal{H}^T} - \frac{1}{\mathcal{H}^{T+1}}\right)e$$

on the domain $e \in [0, (1-\delta)S_{T-1}]$. Importantly, φ is a monotonic function in e. At $e = (1-\delta)S_{T-1}$, $\Delta^T(e) = 0$ by definition, while $\Delta^{T+1}(e) = (1-\delta)\frac{(S_T - S_{T-1})}{\mathcal{H}^{T+1}} > 0$. Hence, $\varphi\left((1-\delta)S_{T-1}\right) > 0$. To show that $\varphi(0) > 0$, which therefore completes the proof, it suffices to show that $\mathcal{H}^T > \mathcal{H}^{T+1}$.

Note that $\mathcal{H}^\tau$ expresses the time-averaged effect of the current partisan lean of policy on the stream of agreements on a foresight of horizon τ. Hence, it is indeed intuitive that the more renegotiations of policy are considered, the more dilute the current partisan-lean of policy becomes on the full stream of agreements. To see this concretely, we can compare $\mathcal{H}^T$ to $\mathcal{H}^{T+1}$ for a particular realization of efficiency shocks $\beta = \{\tilde{\beta}\}$—

$$\mathcal{H}^T(\beta) = 1 + \delta\tilde{\beta}_1 + \ldots + \delta^{\tilde{t}_T}K^{T-1}\left(\prod_{i=1}^{\tilde{t}_T}\tilde{\beta}_i\right)\cdot\mathcal{H}^{T-1}$$

$$= 1 + \delta\tilde{\beta}_1 + \ldots + \left(1 - \frac{\mathcal{H}^{T-1}}{S_{T-2}}\Delta^{T-1}\left(\prod_{i=1}^{\tilde{t}_T}\tilde{\beta}_i\right)\right) \cdot$$
$$\sum_{i=\tilde{t}_T}^{\infty}\left(\delta^i\left[\prod_{j=\tilde{t}_T+1}^{i}\left(1 - \frac{\mathcal{H}^{T-1}}{S_{T-2}}\Delta^{T-1}\left(\prod_{k=\tilde{t}_T+1}^{j}\tilde{\beta}_k\left(\prod_{i=1}^{\tilde{t}_T}\tilde{\beta}_i\right)\right)\right)\tilde{\beta}_j\right]\prod_{i=1}^{\tilde{t}_T-1}\tilde{\beta}_i\right)$$

by substituting $K^{T-1}(e) = \left(1 - \frac{\mathcal{H}^{T-1}}{S_{T-2}}\Delta^{T-1}(e)\right)C^{T-1}(e)$ from the recursive definitions preceding Lemma 4 and further using the derivation of $C^{T-1}(e)$ from (1.29). But now, observe that

$$\Delta^T(e) \cdot \frac{\mathcal{H}^T}{S_{T-1}} = \frac{(1-\delta)S_{T-1} - e}{\mathcal{H}^T} \cdot \frac{\mathcal{H}^T}{S_{T-1}} = 1 - \delta - \frac{e}{S_{T-1}} > 1 - \delta - \frac{e}{S_{T-2}} = \Delta^{T-1}(e) \cdot \frac{\mathcal{H}^{T-1}}{S_{T-2}}.$$

Substituting this inequality into above expression, it then follows that

$$\begin{aligned}
\mathcal{H}^T(\beta) &> 1 + \delta\tilde{\beta}_1 + \ldots \\
&+ \left(1 - \frac{\mathcal{H}^T}{S_{T-1}}\Delta^T\left(\prod_{i=1}^{\tilde{t}_T}\tilde{\beta}_i\right)\right)\sum_{i=\tilde{t}_T}^{\infty}\delta^i\left[\prod_{j=\tilde{t}_T+1}^{i}\left(1 - \frac{\mathcal{H}^T}{S_{T-1}}\Delta^T\left(\prod_{k=\tilde{t}_T+1}^{j}\tilde{\beta}_k \cdot \left(\prod_{i=1}^{\tilde{t}_T}\tilde{\beta}_i\right)\right)\right)\tilde{\beta}_j\right]\left(\prod_{i=1}^{\tilde{t}_T}\tilde{\beta}_i\right) \\
&> 1 + \delta\tilde{\beta}_1 + \ldots \\
&+ \left(1 - \frac{\mathcal{H}^T}{S_{T-1}}\Delta^T\left(\prod_{i=1}^{\tilde{t}_{T+1}}\tilde{\beta}_i\right)\right)\sum_{i=\tilde{t}_{T+1}}^{\infty}\delta^i\left[\prod_{j=\tilde{t}_{T+1}+1}^{i}\left(1 - \frac{\mathcal{H}^T}{S_{T-1}}\Delta^T\left(\prod_{k=\tilde{t}_{T+1}+1}^{j}\tilde{\beta}_k \cdot \left(\prod_{i=1}^{\tilde{t}_{T+1}}\tilde{\beta}_i\right)\right)\right)\tilde{\beta}_j\right] \cdot \prod_{i=1}^{\tilde{t}_{T+1}}\tilde{\beta}_i \\
&= \mathcal{H}^{T+1}(\beta).
\end{aligned}$$

Here, the second inequality results from the fact that because $e^*_{T+1} > e^*_T$, it follows that $\tilde{t}_{T+1} \leq \tilde{t}_T$. That is, with a foresight horizon of $T+1$, the first agreement arrived at in the game will be in force for a weakly shorter duration than with a foresight horizon T, given the same realization of shocks $\beta := \left\{\tilde{\beta}\right\}$. With larger initial agreement duration, by inspection, each element of the sequence $\left(\prod_{i=1}^{j}\tilde{\beta}_j\right)_{j=0}^{\infty}$ is being multiplied by a larger number in the sum above. Taking the expectation over the sequences β gives us $\mathcal{H}^T > \mathcal{H}^{T+1}$. As argued above, this directly implies $\Delta^{T+1}(e) > \Delta^T(e)$, completing the proof of the first part of the proposition.

We now show that $\Delta^T(e)$ is decreasing in δ for arbitrary T. Note that

$$\frac{\partial}{\partial\delta}\Delta^T(e) = \frac{\partial}{\partial\delta}\left[\frac{(1-\delta)S_{T-1} - e}{\mathcal{H}^T}\right] = \frac{\frac{\partial}{\partial\delta}e^*_T \cdot \mathcal{H}^T - \frac{\partial}{\partial\delta}\mathcal{H}^T \cdot (e^*_T - e)}{(\mathcal{H}^T)^2}. \tag{1.30}$$

From Proposition 2, we have shown that $\frac{\partial}{\partial\delta}e^*_T < 0$, so that in order for the expression in (1.30) to be negative, we need only show $\frac{\partial}{\partial\delta}\mathcal{H}^T > 0$. To complete the proof, we demonstrate this auxiliary claim inductively on T. For the base case $T = 1$,

$$\frac{\partial}{\partial\delta}\mathcal{H}^1 = \frac{\partial}{\partial\delta}S_0 = \frac{\overline{\beta}}{(1-\delta\overline{\beta})^2} > 0.$$

Suppose $\frac{\partial}{\partial\delta}\mathcal{H}^T > 0$. We show that this holds for $T+1$ as well. Let $t^* = \min\left\{t : \prod_{j=1}^{t}\tilde{\beta}_j \leq z\right\}$ and define for convenience

$$\hat{K}^T(z) := \mathbb{E}\left[\delta^{t^*} \cdot K^T\left(\prod_{j=1}^{t^*}\tilde{\beta}_j\right)\right]. \tag{1.31}$$

By definition,

$$\mathcal{H}^{T+1} = N(e_*^T) + \hat{K}^T(e_T^*)\mathcal{H}^T.$$

Hence,

$$\begin{aligned}\frac{\partial}{\partial\delta}\mathcal{H}^{T+1} &= \frac{\partial e_*^T}{\partial\delta}\left(N'(e_T^*) + \frac{\partial}{\partial e}\hat{K}^T(e_T^*)\mathcal{H}^T\right) + \frac{\partial}{\partial\delta}\mathcal{H}^T\hat{K}^T(e_T^*) \\ &> \frac{\partial e_*^T}{\partial\delta}\left(N'(e_T^*) + \frac{\partial}{\partial e}\hat{K}^T(e_T^*)\mathcal{H}^T\right),\end{aligned} \tag{1.32}$$

where the inequality follows from the inductive hypothesis. We use again the notation that g_k represents the distribution over any k-prodcut of i.i.d. random variables drawn from the underlying distribution F that governs $\tilde{\beta}$. By definition (1.31),

$$\begin{aligned}\hat{K}^T(z) = \delta\int_0^z K^T(y)g_1(y)\,dy + \delta^2\int_{x=z}^1\int_{y=0}^{\frac{z}{x}} K^T(xy)\cdot g_1(y)g_1(x)\,dy\,dx + \\ \delta^3\int_{x=z}^1\int_{y=0}^{\frac{z}{x}} K^T(xy)\cdot g_1(y)g_2(x)\,dy\,dx + \ldots\end{aligned}$$

By convention, let $g_0 := 0$. Applying the Leibniz rule gives us

$$\begin{aligned}\frac{\partial}{\partial e}\hat{K}^T(z) &= \delta K^T(z)g_1(z) + \sum_{t\geq 1}\delta^{t+1}\left[-g_t(z)\int_0^1 K^T(z\cdot y)g_1(y)\,dy + K^T(z)\int_{x=z}^1\frac{1}{x}g_1\left(\frac{z}{x}\right)g_t(x)\,dx\right] \\ &= \delta K^T(z)g_1(z) + \sum_{t\geq 1}\delta^{t+1}\left[-g_t(z)\int_0^1 K^T(z\cdot y)g_1(y)\,dy + K^T(z)g_{t+1}(z)\right] \\ &= \sum_{t\geq 0}\left\{\delta^{t+1}\left(-g_t(z)\mathbb{E}\left[K^T\left(\tilde{\beta}_{t+1}\cdot z\right)\right] + K^T(z)g_{t+1}(z)\right)\right\},\end{aligned} \tag{1.33}$$

The second equality results from using the definition of the partial distribution function of a product of independent random variables. Utilizing (1.33) and substituting (1.21) for $N'(e_T^*)$ allows us to write

$$N'(e_T^*) + \frac{\partial}{\partial e}\hat{K}^T(e_T^*)\mathcal{H}^T = -e_T^*\cdot\sum_{t\geq 1}\delta^t g_t(e_T^*) + \mathcal{H}^T\cdot\left(K^T(e_T^*) - \delta\mathbb{E}\left[K^T(\tilde{\beta}\cdot e_T^*)\right]\right)\cdot\sum_{t\geq 1}\delta^t g_t(e_T^*), \tag{1.34}$$

where $\tilde{\beta}$ is any shock drawn from the distribution F. By directly substituting in the form of $K^T(e)$,

$$\mathbb{E}\left[K^T(e_T^*) - K^T(\tilde{\beta}\cdot e_T^*)\right] = \sum_{i=0}^{\infty}\delta^i\mathbb{E}\left[\prod_{j=0}^{i}\left(1-\frac{\mathcal{H}^T}{S_{T-1}}\Delta^T\left(\prod_{k=0}^{j}\tilde{\beta}_k\cdot e_T^*\right)\right)\tilde{\beta}_j\right]\frac{e_T^*}{\mathcal{H}^T} - \sum_{i=0}^{\infty}\delta^{i+1}\mathbb{E}\left[\prod_{j=0}^{i}\left(1-\frac{\mathcal{H}^T}{S_{T-1}}\Delta^T\left(\prod_{k=0}^{j}\tilde{\beta}_k\cdot\tilde{\beta}e_T^*\right)\right)\tilde{\beta}_j\cdot\tilde{\beta}\right]\frac{e_T^*}{\mathcal{H}^T} \quad (1.35)$$

Note that for each i, the $(i+1)^{th}$ summand in the former sum is equal to the i^{th} summand in the latter sum of (1.35). Hence, all but the first summand in the former sum remains when we evaluate this difference. We thus obtain

$$\mathbb{E}\left[K^T(e_T^*) - K^T(\tilde{\beta}\cdot e_T^*)\right] = \frac{e_*^T}{\mathcal{H}^T}$$

Substituting back into (1.34) gives

$$N'(e_T^*) + \frac{\partial}{\partial e}\hat{K}^T(e_T^*)\mathcal{H}^T = 0.$$

Therefore, (1.32) straightforwardly gives $\frac{\partial}{\partial\delta}\mathcal{H}^{T+1} > 0$, as desired. □

1.B.2 Proofs for Section 1.4.3

Proposition 6. Let $(L_T)^N(F)$ be the distribution of policies after N agreements have been reached between agents with a foresight horizon T, when the initial policy is drawn from a distribution F. Then for any F, as $N\to\infty$, $(L_T)^N(F)$ converges to G^* symmetric about $\frac{1}{2}$.

Proof. Consider any two distributions F and G on [0,1] from which the initial policy in the sequence of agreements will be drawn.[16] We show that the operator L_T is a contraction mapping. Let $\xi\in[0,e_T^*]\times\{L,R\} := X$ be a particular realization of $(\tilde{e}_T, I_T(\tilde{e}_T))$.

$$\begin{aligned}\int_0^1 |L_T(F)(z) - L_T(G)(z)|\,dz &= \int_0^1\left|\int_X L_T(F)(z|\xi) - L_T(G)(z|\xi)\,dH(\xi)\right|dz\\ &\le \int_0^1\int_X |L_T(F)(z|\xi) - L_T(G|\xi)|\,dH(\xi)\,dz,\\ &\text{by the Cauchy-Schwarz inequality}\\ &= \int_X\int_0^1 |L_T(F)(z|\xi) - L_T(G|\xi)|\,dz\,dH(\xi)\\ &= \int_X\int_0^1 \left|L_T(F|\xi)^{-1}(t) - L_T(G|\xi)^{-1}(t)\right|dt\,dH(\xi), \quad (1.36)\end{aligned}$$

where the final equality in the above series transforms the Riemann-Stieljes integral into the corresponding Lebesgue integral. Here, we integrate over the differences between "percentiles" of the distributions $L_T(F)$ and $L_T(G)$, conditional on a particular realization, $\xi\in X$. Now, note that

[16] With apologies, I am abusing notation here since F and G have been previously been used for distributions corresponding to the decay shocks to the efficiency of policy. Here, they assume no such contextual meaning.

conditional the realization ξ, each percentile of the transformed policy distributions has the same relative ordering as the percentiles of the initial distribution F and G. This is because $a_I^T(e^T, \cdot)$ is a monotone transformation for fixed $\xi = (e^T, I)$. Hence, we have

$$\begin{aligned}\left|L_T(F|\xi)^{-1}(t) - L_T(G|\xi)^{-1}(t)\right| &= \left|a_I^T(e^T, F^{-1}(t)) - a_I^T(e^T, G^{-1}(t))\right|, \text{ where } \xi = (e^T, I) \\ &\leq \lambda \cdot |F^{-1}(t) - G^{-1}(t)|, \text{ from work done in the text.}\end{aligned}$$

Hence, substituting this expression into (1.36), we now have

$$\begin{aligned}\int_0^1 |L_T(F)(z) - L_T(G)(z)| &\leq \int_X \int_0^1 \left|L_T(F|\xi)^{-1}(t) - L_T(G|\xi)^{-1}(t)\right| \, dt \, dH(\xi) \\ &\leq \int_X \int_0^1 \lambda \cdot |F^{-1}(t) - G^{-1}(t)| \, dt \, dH(\xi) \\ &= \lambda \int_0^1 |F^{-1}(t) - G^{-1}(t)| \, dt \\ &= \lambda \int_0^1 |F(z) - G(z)| \, dz,\end{aligned}$$

transforming the Lebesgue integral back into the Riemann-Stieljes integral. We now have that since $\lambda < 1$, L_T is a contraction mapping for any $T \in \mathbb{N}$ over the complete metric space of probability distributions on X, itself a complete metric space. We can therefore apply the Contraction Mapping Theorem to conclude that for any initial distribution F, $(L_T)^N(F)$ converges to a unique distribution G^* as $N \to \infty$. Moreover, note that if we start with F the degenerate distribution that places probability 1 on $\frac{1}{2}$, then at each step N of the process $(L_T)^N(F)$, the resulting distribution will be symmetric about $\frac{1}{2}$. Hence, it must be that the invariant distribution G^* is symmetric about $\frac{1}{2}$. □

1.B.3 Proofs for Section 1.4.4

Proposition 7. Each of the following holds in the unique equilibrium of the legislative bargaining model under Assumptions 1, 2, 3', and 4.

(i) For each $I \in \{L, R\}$, and any T, a_I^T is increasing in $\hat{a}$ (and strictly increasing for $e > 0$).

(ii) For any T and $I \in \{R, L\}$, when $\hat{a} < \frac{1}{2}$, $a_I^T(e, \hat{a}|\gamma)$ is strictly increasing in γ; when $\hat{a} > \frac{1}{2}$, $a_I^T(e, \hat{a}|\gamma)$ strictly decreasing in γ.

(iii) For any T, $W_R^{T+1}(a)$ is strictly increasing in a.

Proof. Our approach will be to show (i)-(iii) using induction on T.

Base Case ($T = 1$).
Note that in this case, since $\mathcal{H}^1$ does not depend on γ, $\Delta^1(e)$ is the same for all γ, and thus for all $\hat{a}$. For (i),

$$\frac{\partial}{\partial \hat{a}} a_L^1(e, \hat{a}|\gamma) = \frac{\partial}{\partial \hat{a}} \left\{ e\hat{a}(1 - \delta\overline{\beta}) + \delta\mathbb{E}\left[a_L^1(\tilde{\beta}_1 e, \hat{a}|\gamma) + \left(1 - G_\gamma\left(a_L^1(\tilde{\beta}_1 e, \hat{a}) + \frac{\Delta^1(\tilde{\beta}_1 e)}{2}\right)\right)\Delta^1(\tilde{\beta}_1 e)\right]\right\}$$

$$= e(1-\delta\overline{\beta}) + \delta\mathbb{E}\left[\frac{\partial}{\partial\hat{a}}a_L^1(\tilde{\beta}_1 e, \hat{a}|\gamma)\cdot\left(1-\Delta^1\left(\tilde{\beta}_1 e\right)\cdot g_\gamma\left(a_L^1(\tilde{\beta}_1 e,\hat{a}|\gamma)+\frac{\Delta^1(\tilde{\beta}_1 e)}{2}\right)\right)\right]. \quad (1.37)$$

Noting that the expression (1.37) for $\frac{\partial}{\partial\hat{a}}a_L^1(e,\hat{a}|\gamma)$ is a function of $\frac{\partial}{\partial\hat{a}}a_L^1(\tilde{\beta}_1 e,\hat{a}|\gamma)$, we can continue expanding the expression (1.37). By recursion, then, we have

$$\frac{\partial}{\partial\hat{a}}a_L^1(e,\hat{a}|\gamma) = e(1-\delta\overline{\beta})\cdot$$
$$\sum_{i=0}^{\infty}\delta^i\mathbb{E}\left[\prod_{j=1}^{i}\left\{\left(1-\Delta^1\left(\prod_{k=1}^{j}\tilde{\beta}_k e\right)\cdot g_\gamma\left(a_L^1\left(\prod_{k=1}^{j}\tilde{\beta}_k e,\hat{a}\middle|\gamma\right)+\frac{\Delta^1\left(\prod_{k=1}^{j}\tilde{\beta}_k e\right)}{2}\right)\right)\cdot\tilde{\beta}_j\right\}\right]. \quad (1.38)$$

Now, denote for each $i,\ell\geq 0$

$$s_i := \mathbb{E}\left[\prod_{j=1}^{i}\left\{\left(1-\Delta^1\left(\prod_{k=1}^{j}\tilde{\beta}_k e\right)\cdot g_\gamma\left(a_L^1\left(\prod_{k=1}^{j}\tilde{\beta}_k e,\hat{a}\middle|\gamma\right)+\frac{\Delta^1\left(\prod_{k=1}^{j}\tilde{\beta}_k e\right)}{2}\right)\right)\cdot\tilde{\beta}_j\right\}\right]$$
$$\mathcal{S}_\ell := \sum_{i=0}^{\ell}\delta^i s_i,$$

where s_i is the i^{th} summand of the sum in (1.38) (divided by δ^i) and $\mathcal{S}_\ell$ the partial sum of the first $\ell+1$ summands in (1.38), respectively (noting the first summand has index $i=0$). For referential purposes, we denote the remaining argument for (i) as ($\heartsuit$) .

Trivially, $\mathcal{S}_0 = s_0 = 1 > 0$. Call then by $i^*\geq 1$ the smallest-index i such that the summand $s_i < 0$. If $i^*=\infty$, each summand is nonnegative after the first and therefore the expression (1.38) is strictly positive. So, consider i^* finite. Note that because $g_\gamma(x)\leq 2$ for any $x\in[0,1]$, it follows that

$$s_{i^*} > \left(1-2\mathbb{E}\left[\Delta^1\left(\prod_{j=1}^{i}\tilde{\beta}_j e\right)\right]\right)\cdot s_{i^*-1} > (1-2(1-\delta))\cdot s_{i^*-1}.$$

We thus have

$$\begin{aligned}\mathcal{S}_{i^*} = \mathcal{S}_{i^*-1}+\delta^{i^*}s_{i^*} &> \mathcal{S}_{i^*-1}+\delta^k s_{i^*-1}\left(1-2(1-\delta)\right)\\ &= \mathcal{S}_{i^*-2}+\delta^{i^*-1}s_{i^*-1}\left(1+\delta(1-2(1-\delta))\right),\end{aligned}$$

where $\mathcal{S}_{-1}=0$. $\mathcal{S}_{i^*-2}\geq 0$, and by hypothesis on i^*, $s_{i^*-1}\geq 0$, with the inequality strict if $i^*=1$. The sum $1+\delta(1-2(1-\delta))>0$ so that $\mathcal{S}_{i^*}>0$. In fact, recognizing the multiplicative, cumulative nature of the summands s_i, we can use the same approach to bound any $\mathcal{S}_k$ for $k\geq i^*$

$$\begin{aligned}\mathcal{S}_k > \mathcal{S}_{k-1}+\delta^k\frac{s_k}{1-\delta} &> \mathcal{S}_{k-1}+s_{k-1}\cdot\delta^k\frac{(2\delta-1)}{1-\delta}\\ &> \mathcal{S}_{k-1}-\delta^k s_{k-1} > 0,\end{aligned}$$

where the penultimate inequality results from the fact that $\frac{(2\delta-1)}{1-\delta} > -1$. Hence, the expression (1.38) must also be strictly positive, completing the proof of (i) in the base case.

We show (ii) for the case where $\hat{a} < \frac{1}{2}$ since the opposite case is analogous. We require the following lemma—

Lemma 5. *Suppose $W_R^T(a)$ is increasing in a and $a_I^T(e, \hat{a}|\gamma)$ is increasing in $\hat{a}$ for $I \in \{R, L\}$. If $\hat{a} < \frac{1}{2}$, then $\mu^T(e, \hat{a}|\gamma)$ is strictly decreasing in e; if $\hat{a} > \frac{1}{2}$, then $\mu^T(e, \hat{a}|\gamma)$ is strictly increasing in e.*

Proof (of Lemma 1.B.3). Observe that since $W_R^T(a)$ is increasing, the hypothesis of Lemma 2 is satisfied. Therefore,

$$\begin{aligned}
\mu^T(e, \hat{a}|\gamma) &= \frac{W_R^T(a_L^T(e, \hat{a}|\gamma)) + W_R^T(a_R^T(e, \hat{a}|\gamma)) - [W_R^T(a_L^T(e, \hat{a}|\gamma)) + W_R^T(a_L^T(e, \hat{a}|\gamma))]}{2S_{T-1}} \\
&= \frac{2W_R^T(a_L^T(e, \hat{a}|\gamma)) + (1-\delta)S_{T-1} - e}{2S_{T-1}} \\
&= \frac{W_R^T(a_L^T(e, \hat{a})) + (1-\delta)S_{T-1}/2 - e/2}{S_{T-1}}.
\end{aligned}$$

Using this expression,

$$\begin{aligned}
\frac{\partial}{\partial e}\mu^T(e, \hat{a}|\gamma) &= \frac{1}{S_{T-1}} \cdot \frac{\partial}{\partial e}\left\{ W_R^T(a_L^T(e, \hat{a})) + \frac{(1-\delta)S_{T-1}}{2} - \frac{e}{2} \right\} \\
&= \frac{1}{S_{T-1}}\Bigg(-\frac{1}{2} + \frac{\partial}{\partial e}\Bigg\{ e\hat{a} + \delta\mathbb{E}\Big[W_R^T(a_L^T(\tilde{\beta}_1 e, \hat{a})) + \\
&\qquad (1 - G_\gamma(\mu^T(\tilde{\beta}_1 e, \hat{a}|\gamma)))((1-\delta)S_{T-1} - \tilde{\beta}_1 e) \Big] \Bigg\} \Bigg) \\
&= \frac{1}{S_{T-1}}\Bigg(\left(\hat{a} - \frac{1}{2}\right) + \delta\mathbb{E}\Big[-g_\gamma(\mu^T(\tilde{\beta}_1 e, \hat{a}|\gamma)) \cdot ((1-\delta)S_{T-1} - \tilde{\beta}_1 e) \cdot \frac{\partial}{\partial e}\left\{ \mu^T(\tilde{\beta}_1 e, \hat{a}|\gamma) \right\} \\
&\qquad - (1 - G_\gamma(\mu^T(\tilde{\beta}_1 e, \hat{a}|\gamma)))\tilde{\beta}_1 + \frac{\partial}{\partial e}\left\{ W_R^T(a_L^T(\tilde{\beta}_1 e, \hat{a})) \right\} \Big] \Bigg) \\
&= \frac{1}{S_{T-1}} \cdot \mathbb{E}\left[\left(\hat{a} - \frac{1}{2}\right) + \delta\left(G_\gamma(\mu^T(\tilde{\beta}_1 e, \hat{a}|\gamma)) - \frac{1}{2} \right)\tilde{\beta}_1 \right] + \\
&\qquad \delta\mathbb{E}\left[\left(1 - g_\gamma(\mu^T(\tilde{\beta}_1 e, \hat{a}|\gamma)) \cdot ((1-\delta)S_{T-1} - \tilde{\beta}_1 e)\right) \cdot \frac{\partial}{\partial e}\left\{ \mu^T(\tilde{\beta}_1 e, \hat{a}|\gamma) \right\} \right],
\end{aligned}$$

where the final equality results from rewriting $W_R^T(a_L^T(\tilde{\beta}_1 e, \hat{a}))$ in terms of $\mu^T(\tilde{\beta}_1 e, \hat{a}|\gamma)$. Note that this is a recursive equation, whose solution is given by

$$\begin{aligned}
\frac{\partial}{\partial e}\mu^T(e, \hat{a}|\gamma) = \frac{1}{S_{T-1}} \sum_{i=0}^{\infty} \delta^i \mathbb{E}\Bigg[\prod_{j=1}^{i} \left\{ \left(1 - g_\gamma\left(\mu^T\left(\prod_{k=1}^{j} \tilde{\beta}_k \cdot e, \hat{a}|\gamma \right) \right) \cdot \frac{(1-\delta)S_{T-1} - \prod_{k=1}^{j} \tilde{\beta}_k \cdot e}{S_{T-1}} \right) \tilde{\beta}_j \right\} \cdot \\
\left(\hat{a} - \frac{1}{2} + \delta\left(G_\gamma\left(\mu^T\left(\prod_{k=1}^{i+1} \tilde{\beta}_k \cdot e, \hat{a}|\gamma \right) \right) - \frac{1}{2} \right) \tilde{\beta}_{i+1} \right) \Bigg]. \quad (1.39)
\end{aligned}$$

When $\hat{a} = \frac{1}{2}$, the game is fully symmetric so that for any e' at which agreement occurs, $\mu^T(e', \hat{a}|\gamma) = \frac{1}{2}$ and therefore $G_\gamma(\mu^T(e', \hat{a}|\gamma)) = \frac{1}{2}$. Consider first $\hat{a} < \frac{1}{2}$. Here, by the fact that both W_R^T and a_L^T are increasing, $G_\gamma\left(\mu^T\left(\prod_{k=1}^{i+1}\tilde{\beta}_k \cdot e, \hat{a}|\gamma\right)\right) < \frac{1}{2}$, for each i. From (1.39),

$$\frac{\partial}{\partial e}\mu^T(e, \hat{a}|\gamma) < \frac{\hat{a} - \frac{1}{2}}{S_{T-1}} \sum_{i=0}^{\infty} \delta^i \mathbb{E}\left[\prod_{j=1}^{i}\left\{\left(1 - g_\gamma\left(\mu^T\left(\prod_{k=1}^{j}\tilde{\beta}_k \cdot e, \hat{a}|\gamma\right)\right)\right) \cdot \frac{(1-\delta)S_{T-1} - \prod_{k=1}^{j}\tilde{\beta}_k \cdot e}{S_{T-1}}\right)\tilde{\beta}_j\right\}\right].$$

Finally, replicating argument (♡) shows that the sum in the above equation must be positive. Thus, for $\hat{a} < \frac{1}{2}$,

$$\frac{\partial}{\partial e}\mu^T(e, \hat{a}|\gamma) < 0,$$

as desired. The case of $\hat{a} > \frac{1}{2}$ is analogous noting that the quantities $\hat{a} - \frac{1}{2}$ and

$$G_\gamma\left(\mu^T\left(\prod_{k=1}^{i+1}\tilde{\beta}_k \cdot e, \hat{a}|\gamma\right)\right) - \frac{1}{2}$$

will both become positive. This completes the proof of the lemma. □

We can now prove (ii) for the base case $T = 1$. Define for any function $f : [0,1]^2 \to [0,1]$,

$$Q_{\hat{a}}(f, e) := \frac{1}{2}f(e, \hat{a})\left(\frac{1}{2} - f(e, \hat{a})\right).$$

We then have

$$\begin{aligned}\frac{\partial}{\partial\gamma}a_L^1(e, \hat{a}|\gamma) &= \delta\mathbb{E}\left[\frac{\partial}{\partial\gamma}\left\{a_L^1(\tilde{\beta}_1 e, \hat{a}|\gamma) + \left(1 - G_\gamma\left(\mu^1(\tilde{\beta}_1 e, \hat{a}|\gamma)\right)\right)\Delta^1(\tilde{\beta}_1 e)\right\}\right] \\ &= \delta\mathbb{E}\left[\Delta^1(\tilde{\beta}_1 e) \cdot Q_{\hat{a}}\left(\mu^1(\cdot|\gamma), \tilde{\beta}_1 e\right)\right] + \\ &\qquad \delta\mathbb{E}\left[\left(1 - \left(1 + \gamma\left(\mu^1(\tilde{\beta}_1 e, \hat{a}|\gamma) - \frac{1}{4}\right)\right)\Delta^1(\tilde{\beta}_1 e)\right)\frac{\partial}{\partial\gamma}a_L^1(\tilde{\beta}_1 e, \hat{a}|\gamma)\right].\end{aligned}$$

Solving the above recursive equation gives us

$$\frac{\partial}{\partial\gamma}a_L^1(e, \hat{a}|\gamma) = \sum_{i=1}^{\infty}\delta^i\mathbb{E}\left[\prod_{j=1}^{i-1}\hat{r}_j^1 \cdot Q_{\hat{a}}\left(\mu^1(\cdot|\gamma), \prod_{k=1}^{i}\tilde{\beta}_k \cdot e\right)\Delta^1\left(\prod_{k=1}^{i}\tilde{\beta}_k \cdot e\right)\right], \text{ where} \quad (1.40)$$

$$\hat{r}_j^1 := 1 - \left(1 + \gamma\left(\mu^1(\prod_{k=1}^{j}\tilde{\beta}_k \cdot e, \hat{a}|\gamma) - \frac{1}{4}\right)\right)\Delta^1\left(\prod_{k=1}^{j}\tilde{\beta}_k \cdot e\right).$$

Let also the i^{th} summand of 1.40 be denoted r_i and the sum of the first k summands $\mathcal{R}_k$, with $\mathcal{R}_0 = 0$ by convention. For referential purposes, we call the remainder of the argument completing the proof of (ii) in the base case (♠). Note that $r_1 > 0$. If each $r_i > 0$, then the right hand side of (1.40) is positive and the proof is complete. Hence, suppose there exists some minimal $i^* \geq 2$ such

that $r_{i^*} < 0$. . Then

$$\mathcal{R}_{i^*} = \mathcal{R}_{i^*-2} + r_{i^*-1}\left(1 + \delta\hat{r}^1_{i^*-1}\frac{Q_{\hat{a}}\left(\mu^1(\cdot|\gamma), \prod_{j=1}^{i^*}\tilde{\beta}_j \cdot e\right)\Delta^1\left(\prod_{j=1}^{i^*}\tilde{\beta}_j \cdot e\right)}{Q_{\hat{a}}\left(\mu^1(\cdot|\gamma), \prod_{j=1}^{i^*-1}\tilde{\beta}_j \cdot e\right)\Delta^1\left(\prod_{j=1}^{i^*-1}\tilde{\beta}_j \cdot e\right)}\right), \tag{1.41}$$

and $\mathcal{R}_{i^*-2} \geq 0$, $r_{i^*-1} > 0$, by minimality of i^*. In order for $r_{i^*} < 0$, it must be that $\hat{r}^1_{i^*-1} < 0$, which in turn implies that $\mu^1(\prod_{k=1}^{i^*-1}\tilde{\beta}_k \cdot e, \hat{a}|\gamma) > \frac{1}{4}$. Recall that since $\gamma \in [-4, 4]$, it follows that

$$\hat{r}^1_{i^*-1} > 1 - 3\Delta^1\left(\prod_{j=1}^{i^*-1}\tilde{\beta}_j \cdot e\right).$$

In fact, hypothesis, since $\hat{r}^1_{i^*-1} < 0$,

$$\Delta^1\left(\prod_{j=1}^{i^*-1}\tilde{\beta}_j \cdot e\right) > \left[1 + \gamma\left(\mu^1(\prod_{k=1}^{j}\tilde{\beta}_k \cdot e, \hat{a}|\gamma) - \frac{1}{4}\right)\right]^{-1},$$

which gives $\delta \leq \frac{2}{3}$. Hence,

$$\delta\frac{\hat{r}^1_{i^*-1}}{\Delta^1\left(\prod_{j=1}^{i^*-1}\tilde{\beta}_j \cdot e\right)} > \frac{\delta}{1-\delta} - 3\delta \geq -1,$$

for all $\delta \leq \frac{2}{3}$. Therefore, from (1.41), we have

$$\mathcal{R}_{i^*} > \mathcal{R}_{i^*-2} + r_{i^*-1}\left(1 - \frac{Q_{\hat{a}}\left(\mu^1(\cdot|\gamma), \prod_{j=1}^{i^*}\tilde{\beta}_j \cdot e\right)}{Q_{\hat{a}}\left(\mu^1(\cdot|\gamma), \prod_{j=1}^{i^*-1}\tilde{\beta}_j \cdot e\right)}\right).$$

Because $Q_{\hat{a}}(\mu^1(\cdot|\gamma), e')$ *decreases* as $\mu^1(e', \hat{a}|\gamma) \in \left(\frac{1}{4}, \frac{1}{2}\right)$ *increases*, using Lemma 1.B.3, we then see that

$$Q_{\hat{a}}\left(\mu^1(\cdot|\gamma), \prod_{j=1}^{i^*-1}\tilde{\beta}_j \cdot e\right) > Q_{\hat{a}}\left(\mu^1(\cdot|\gamma), \prod_{j=1}^{i^*}\tilde{\beta}_j \cdot e\right).$$

Hence,

$$\mathcal{R}_{i^*} > \mathcal{R}_{i^*-2} \geq 0.$$

Finally, by Lemma 1.B.3, $\mu^1(e', \hat{a}|\gamma)$ is decreasing in e' so that $\hat{r}^1_j < 0$ for each $j \geq i^* - 1$. Hence, once $r_{i^*} < 0$, the sign of the terms r_i will alternate for $i \in \{i^*, i^* + 1, \ldots\}$. The above argument thus naturally generalizes to show that each partial sum $\mathcal{R}_{i^*+2k} > \mathcal{R}_{i^*+2(k-1)} > 0$. This immediately gives us that the right-hand side of (1.40) is positive, which shows that when $\hat{a} < \frac{1}{2}$, $a^1_L(e, \hat{a}|\gamma)$ is increasing in γ. Note that since $\Delta^1(e)$ does not depend on γ or $\hat{a}$, the same holds for $a^1_R(e, \hat{a}|\gamma)$ as well. Additionally, the value funtions $W^T_R(a^1_I(e, \hat{a}|\gamma))$ are trivially increasing in γ since $W^1_R(a) = \frac{a}{1-\delta\bar{\beta}}$. The case of $\hat{a} > \frac{1}{2}$ is once again, analogous.

Note that $a_L^1(e,\hat{a}|\gamma)$ and $a_R^1(e,\hat{a}|\gamma)$ are increasing in γ while their midpoint does not exceed $\frac{1}{2}$ for any γ by the fact that $\hat{a} < \frac{1}{2}$ and the policy functions are strictly increasing in $\hat{a}$ from part (i). Thus, their midpoint is strictly increasing in γ. This completes the proof of (ii) in the base case.

For (iii),

$$W_R^2(a) = a \cdot \mathbb{E}[\tilde{\eta}^2] + \mathbb{E}\left[\delta^{\tilde{t}_2}\tilde{\eta}^1 \cdot a^1(\tilde{e}^1, a|\gamma)\right]$$
$$\implies \frac{\partial}{\partial a}W_R^2(a) = \mathbb{E}[\tilde{\eta}^2] + \mathbb{E}\left[\delta^{\tilde{t}_2}\tilde{\eta}^1 \cdot \frac{\partial}{\partial a}a_L^1(\tilde{e}^1,\hat{a})\left(1 - g_\gamma\left(a_L^1(\tilde{e}^1,a) + \frac{\Delta^1(\tilde{e}^1)}{2}\right)\Delta^1(\tilde{e}^1)\right)\right]$$

Because $\mathbb{E}[\tilde{\eta}^2] \geq 1$; $\frac{\partial}{\partial a}a_L^1(\tilde{e}^1,\hat{a}) < 1$, by direct inspection of (1.38); and

$$1 - g_\gamma\left(a_L^1(\tilde{e}^1,a) + \frac{\Delta^1(\tilde{e}^1)}{2}\right)\Delta^1(\tilde{e}^1) > -(1-\delta),$$

from previous work, the above expression must be positive. This completes the proof of the base case.

Inductive Step.
Suppose statements (i)-(iii) hold for $T-1$. We want to show these also hold for T. First,

$$\begin{aligned} W_R^T(a_L^T(e,\hat{a}|\gamma)) &= e\hat{a} + \delta\mathbb{E}\left[W_R^T(a^T(\tilde{\beta}_1 e,\hat{a}|\gamma))\right] \\ &= e\hat{a} + \delta\mathbb{E}\Bigg[G_\gamma\left(\mu^T(\tilde{\beta}_1 e,\hat{a}|\gamma)\right)W_R^T(a_L^T(\tilde{\beta}_1 e,\hat{a}|\gamma)) \\ &\quad + \left(1 - G_\gamma\left(\mu^T(\tilde{\beta}_1 e,\hat{a}|\gamma)\right)\right)W_R^T(a_R^T(\tilde{\beta}_1 e,\hat{a}|\gamma))\Bigg], \end{aligned} \tag{1.42}$$

$$W_L^T(a_R^T(e,\hat{a}|\gamma)) = S_{T-1} - W_R^T(a_R^T(e,\hat{a}|\gamma)) = e(1-\hat{a}) + \delta\mathbb{E}\left[S_{T-1} - W_R^T(a^T(\tilde{\beta}_1 e,\hat{a}|\gamma))\right]. \tag{1.43}$$

For (i), begin by noting from (1.43) that for any $e \leqq e_T^*$ and $I \in \{R, L\}$,

$$\frac{\partial}{\partial\hat{a}}W_R^T(a_I^T(e,\hat{a})) = e + \delta\mathbb{E}\left[\frac{\partial}{\partial\hat{a}}\left\{W_R^T(a^T(\tilde{\beta}_1 e,\hat{a}|\gamma))\right\}\right]$$

Using this identity, taking the derivative of (1.42) with respect to $\hat{a}$ gives

$$\frac{\partial}{\partial\hat{a}}\left\{W_R^T(a_L^T(e,\hat{a}|\gamma))\right\} = e + \delta\mathbb{E}\left[\frac{\partial}{\partial\hat{a}}\left\{W_R^T(a_R^T(\tilde{\beta}_1 e,\hat{a}|\gamma))\right\}\left(1 - g_\gamma\left(\mu^T(\tilde{\beta}_1 e,\hat{a}|\gamma)\right)\left((1-\delta) - \frac{\tilde{\beta}_1 e}{S_{T-1}}\right)\right)\right],$$

which is a recursive equation as before. Solving gives us

$$\frac{\partial}{\partial\hat{a}}\left\{W_R^T(a_L^T(e,\hat{a}|\gamma))\right\} = e\sum_{i=0}^{\infty}\delta^i \cdot \mathbb{E}\left[\prod_{j=1}^{i}\left(1 - g_\gamma\left(\mu^T(\prod_{k=1}^{j}\tilde{\beta}_k \cdot e,\hat{a}|\gamma)\right) \cdot \left((1-\delta) - \frac{\prod_{k=1}^{j}\tilde{\beta}_k \cdot e}{S_{T-1}}\right)\right)\tilde{\beta}_j\right]. \tag{1.44}$$

Here, (1.44) is exactly the T- analogue of (1.38), which was derived for the base case. Thus, repeating the argument ($\heartsuit$) for (1.44) shows that $\frac{\partial}{\partial \hat{a}}\left\{W_R^T(a_L^T(e,\hat{a}|\gamma))\right\} > 0$. Finally, by inductive hypothesis $\frac{\partial}{\partial a}W_R^T(a) > 0$ so that

$$\frac{\partial}{\partial \hat{a}}\left\{W_R^T(a_L^T(e,\hat{a}|\gamma))\right\} = \frac{\partial}{\partial a}W_R^T\Big|_{a=a_L^T(e,\hat{a}|\gamma)} \cdot \frac{\partial}{\partial \hat{a}}a_L^T(e,\hat{a}|\gamma) > 0$$

immediately implies that $\frac{\partial}{\partial \hat{a}}a_L^T(e,\hat{a}|\gamma) > 0$. The proof for $a_R^T(e,\hat{a}|\gamma)$ is exactly analogous, showing (i).

Next, for (ii), supposing that $\hat{a} < \frac{1}{2}$,

$$\begin{aligned}
\frac{\partial}{\partial \gamma}\left\{W_R^T(a_L^T(e,\hat{a}|\gamma))\right\} &= \frac{\partial}{\partial \gamma}\left\{e\hat{a} + \delta\mathbb{E}\left[W_R^T(a_L^T(\tilde{\beta}_1 e,\hat{a}|\gamma)) + (1-G_\gamma(\mu^T(\tilde{\beta}_1 e,\hat{a}))\cdot((1-\delta)S_{T-1} - \tilde{\beta}_1 e)\right]\right\} \\
&= \delta\mathbb{E}\Bigg[Q_{\hat{a}}\left(\mu^T(\cdot|\gamma),\tilde{\beta}_1 e\right)((1-\delta)S_{T-1} - \tilde{\beta}_1 e) + \\
&\qquad \left(1 - \frac{g_\gamma(\mu^T(\tilde{\beta}_1 e,\hat{a}|\gamma))}{S_{T-1}}((1-\delta)S_{T-1} - \tilde{\beta}_1 e)\right)\cdot\frac{\partial}{\partial\gamma}\left\{W_R^T(a_L^T(\tilde{\beta}_1 e,\hat{a}|\gamma)\right\}\Bigg] \\
&= \delta\mathbb{E}\left[Q_{\hat{a}}\left(\mu^T(\cdot|\gamma),\tilde{\beta}_1 e\right)((1-\delta)S_{T-1} - \tilde{\beta}_1 e)\right] \\
&\quad + \delta\mathbb{E}\Bigg[\left(1 - \left(1+\gamma\left(\mu^T(\tilde{\beta}_1 e,\hat{a}|\gamma) - \frac{1}{4}\right)\right)\left(1 - \delta - \frac{\tilde{\beta}_1 e}{S_{T-1}}\right)\right)\cdot \\
&\quad \frac{\partial}{\partial\gamma}\left\{W_R^T(a_L^T(\tilde{\beta}_1 e,\hat{a}|\gamma)\right\}\Bigg],
\end{aligned}$$

which is again a recursive equation with solution given by

$$\frac{\partial}{\partial\gamma}\left\{W_R^T(a_L^T(e,\hat{a}|\gamma))\right\} = S_{T-1}\cdot\sum_{i=1}^{\infty}\delta^i\mathbb{E}\left[\prod_{j=1}^{i-1}\hat{r}_j^T\cdot Q_{\hat{a}}\left(\mu^T(\cdot|\gamma),\prod_{k=1}^{i}\tilde{\beta}_k\cdot e\right)\cdot\left(1-\delta-\frac{\prod_{k=1}^{i}\tilde{\beta}_k\cdot e}{S_{T-1}}\right)\right], \tag{1.45}$$

where

$$\hat{r}_j^T := 1 - \left(1+\gamma\left(\mu^T(\prod_{k=1}^{j}\tilde{\beta}_k\cdot e,\hat{a}|\gamma) - \frac{1}{4}\right)\right)\left(1-\delta-\frac{\prod_{k=1}^{j}\tilde{\beta}_k\cdot e}{S_{T-1}}\right).$$

Of course, (1.45) is exactly the T-analogue of (1.40) so that by replicating argument ($\spadesuit$) from the base case we complete the proof of (ii). Again, the case of $\hat{a} > \frac{1}{2}$ and the proof for $a_R^T(e,\hat{a}|\gamma)$ are exactly analogous.

Finally, for (iii),

$$\begin{aligned}
W_R^{T+1}(a) &= a\cdot\mathbb{E}[\tilde{\eta}^{T+1}] + \mathbb{E}\left[\delta^{\tilde{t}_{T+1}}W_R^T(a^T(\tilde{e}^T,a|\gamma))\right] \\
\implies \frac{\partial}{\partial a}W_R^{T+1}(a) &= \mathbb{E}[\tilde{\eta}^{T+1}] + \mathbb{E}\left[\delta^{\tilde{t}_{T+1}}\cdot\frac{\partial}{\partial a}\left\{W_R^T(a_L^T(\tilde{e}^T,a)) + (1-G_\gamma(\mu^T(\tilde{e}^T,a|\gamma)))\cdot((1-\delta)S_{T-1} - \tilde{e}^T)\right\}\right] \\
&= \mathbb{E}[\tilde{\eta}^{T+1}] + \\
&\quad \mathbb{E}\left[\delta^{\tilde{t}_{T+1}}\left(\frac{\partial}{\partial a}\left\{W_R^T(a_L^T(\tilde{e}^T,a))\right\} - g_\gamma(\mu^T(\tilde{e}^T,a|\gamma))\cdot\frac{\partial\mu^T(\tilde{e}^T,a|\gamma)}{\partial a}((1-\delta)S_{T-1} - \tilde{e}^T)\right)\right]
\end{aligned}$$

$$= \mathbb{E}[\tilde{\eta}^{T+1}] + \mathbb{E}\left[\delta^{\tilde{t}_{T+1}} \cdot \frac{\partial}{\partial a}\left\{W_R^T(a_L^T(\tilde{e}^T, a))\right\}\left(1 - g_\gamma(\mu^T(\tilde{e}^T, a|\gamma)) \cdot \left((1-\delta) - \frac{\tilde{e}^T}{S_{T-1}}\right)\right)\right].$$

Note that—

(1) $\mathbb{E}[\tilde{\eta}^{T+1}] \geq 1$

(2) $\frac{\partial}{\partial a} W_R^T(a_L^T(\tilde{e}^T, a)) < \frac{1}{1-\delta}$ by direct inspection of (1.44).

(3) From previous work within argument $(\heartsuit)$, $1 - g_\gamma(\mu^T(\tilde{e}^T, a|\gamma)) \cdot \left((1-\delta) - \frac{\tilde{e}^T}{S_{T-1}}\right) > -(1-\delta)$

Putting facts (1) - (3) together gives us that $\frac{\partial}{\partial a} W_R^{T+1}(a) > 0$. This completes the proof of the inductive step and thus the proof of the proposition. □

1.C A Micro-Foundation for Limited Foresight

As discussed in Section 1.2, negotiators exhibit *limited foresight* when they give import to only a finite number of downstream agreements that result from the current implementation of policy. One canonical way in which this behavior arises is through the *inability* of agents to fully account for the impact the current policy implementation will have on optimal choices at future decision nodes. Agents then respond by truncating the dynamic linkage of downstream agreements to a chain of length T.

1.C.1 Endogenizing Foresight in Equilibrium

To illustrate how this may happen, suppose the status quo is $(e, \hat{a})$ with incumbent I. For a foresight horizon T, the proposal functions of the parties $\{a_i^\tau\}_{\tau=1}^T$ are well-defined at those states at which agreement is reached, within any equilibrium (with slight abuse of notation, for simplicity, I drop σ here to denote a particular equilibrium in the general setting).

(1) The agreement reached in the current state has only one possibility— $a_I^T(e, \hat{a})$— given the information set of the agents.

(2) The subsequent agreement that is arrived at on the projected path of play of the agents, on the other hand, is uncertain. This is given by $a_{I_{T-1}}^{T-1}(\tilde{e}^{T-1}, a_I^T(e, \hat{a}))$, where the uncertainty is over both the eventual incumbent I_{T-1} when this agreement is hatched, as well as the efficiency level, $\tilde{e}^{T-1}$ at which this occurs.

(3) Continuing along, the next agreement arrived at is $a_{I_{T-2}}^{T-2}\left(\tilde{e}^{T-2}, a_{I_{T-1}}^{T-1}(\tilde{e}^{T-1}, a_I^T(e, \hat{a}))\right)$. Here, there is uncertainty on *four dimensions*: the incumbent I_{T-2} and the efficiency level $\tilde{e}^{T-2}$ when this agreement is made, along with the incentives induced by the *previous agreement*, $a_{I_{T-1}}^{T-1}$, which itself encapsulated uncertainty along two dimensions.

Thus, there are even more possibilities as one projects agreements further downstream from the current negotiation. In particular, the uncertainty inherent in one negotiation continues to compound

when considering all agreements whose outcomes are dependent on the specific nature of its resolution.

Formally, given a foresight horizon T, and initial state $((e,\hat{a}),I)$, we let $X^{(T+1)-\tau}((e,\hat{a}),I)$ be the set of possible agreements reached τ re-negotiations from the present. That is, $a \in X^{(T+1)-\tau}((e,\hat{a}),I)$ if and only if there exists some sequence of $\{\tilde{e}^{\tau'}\}_{\tau'=T}^{T-\tau}$ and $\{I_{\tau'}\}_{\tau'=T}^{T-\tau}$ such that $a = a_{I_{T-\tau}}^{T-\tau}\left(\tilde{e}^{T-\tau}, a_{I_{T-\tau+1}}^{T-\tau+1}\left(\tilde{e}^{T-\tau+1},\ldots\right)\right)$ in any given equilibrium σ. We can now define the following notion of total strategic complexity in the problem faced by a decision-maker by foresight horizon T:

Definition 9. *For given equilibrium σ, the total statistical entropy over future agreements within a foresight horizon T is denoted by*

$$\mathcal{E}^T((e,\hat{a}),I) := -\sum_{\tau=1}^{T}\left(\sum_{x\in X^{(T+1)-\tau}} \mathbb{P}(x)\cdot\ln\mathbb{P}(x)\right). \tag{1.46}$$

This definition, of course, is analogous to the standard definition of entropy in an information theory setting from Shannon (1948). Here, Shannon entropy is summed over each of the agreements up to T links from the present decision node. Since the sets $X^{(T+1)-\tau}$ themselves depend on the equilibrium σ, $\mathcal{E}^T((e,\hat{a},I)$ is itself an equilibrium object.

We introduce also a *cognitive constraint* $B \in (0,1)$ corresponding to the agents that represents the degree of compounded uncertainty they are able to accurately incorporate within their maximization problem. Requiring that in equilibrium, $\mathcal{E}^T((e,\hat{a},I)$ not exceed B thus captures the generic inability of the negotiator to consider the infinite chain of agreements that flow from the current one. This requirement is delineated in the following refinement of our earlier notion of equilibrium, which endogenizes the foresight horizon using precisely the above reasoning.

Definition 10 (Endogenous Foresight Equilibrium). *Given a cognitive constraint B, exponential discount factor δ, and stochastic decay shock $\tilde{\beta}$, an endogenous foresight equilibrium of the game is given by strategies σ_i^T (as given in Definition 3) for $i \in \{R,L\}$ and a foresight horizon T such that*

(1) The strategy profile $\sigma := (\sigma_L^T, \sigma_R^T)$ is a Markov perfect equilibrium

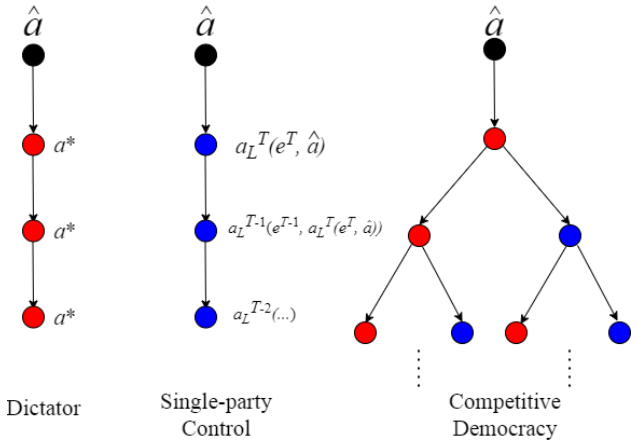
(2) The foresight horizon T is the maximal T' such that for each $(e,\hat{a})$ at which agreement occurs under σ and every incumbent I,

$$\mathcal{E}^{T'}((e,\hat{a}),I) \leq B$$

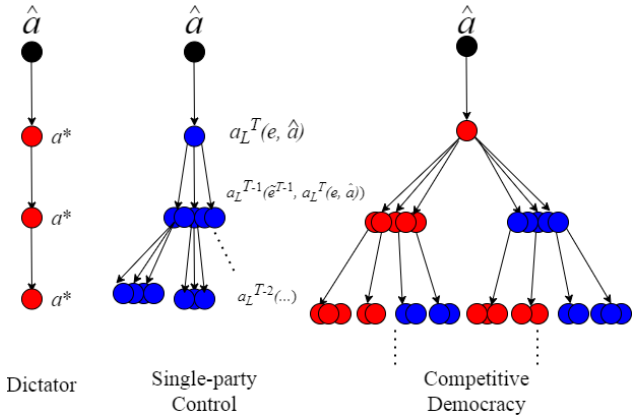
Finally, note that the uniqueness of equilibrium we derived in Proposition 4 continues to hold under this alternative definition. This is because, since there is a unique equilibrium for each foresight horizon T, the unique endogenous foresight equilibrium equilibrium is given from the mapping from $B \to T$ that is induced by (2) in Definition 10.

1.C.2 Discussion

One interesting thing to note is that under this endogenization of the foresight horizon, the strategic setting of the negotiations themselves can have a large impact on the foresight horizon of agents, and thus the outcome.



(a)



(b)

Figure 1.9: Total statistical entropy increases due to uncertainty over both the incumbent at the time of a break-through in negotiations along with that over the efficiency of policy at any future of time.

For illustration, consider legislative bargaining. Various situations are shown in Figure 1.9 above. The bargaining game can be trivial, in the sense that decisions are made by a dictator D in each period; here, there is never any strategic complexity, since in each period, the dictator simply implements her ideal point, a^*. Given this behavior, it easy to see that for any $T \in \mathbb{N}$, and from any status quo policy $\hat{p}$, $\mathcal{E}^T((e, \hat{a}), D) = 0$.

We might also consider a democracy in which there is single-party control over legislative institutions; here, the opposition does not have an effective means to oppose legislation. This situation mirrors the United States Congress in the 1960's and early 1970's when Democrats controlled a *filibuster-proof* majority in the Senate along with a majority in the House of Representatives. If the shocks $\tilde{\beta}$ were deterministic, then there would be no uncertainty within the entire chain of agreements passed since agents are certain about the incentives— the status quo policy and the identity of the incumbent— they will face in each of the relevant decision nodes. Hence, uncertainty only exists through the stochasticity of the shock distributions $\tilde{\beta}$.

Finally, in a competitive democracy, such as the one that has existed in the United States from 1994 (when Republicans gained control of the House for the first time in five decades) until the present, the entropy is largest. Here, uncertainty accumulates not only through uncertainty over

the identity of the incumbent, but also through uncertainty over the incentives that any incumbent would face when $\tilde{\beta}$ is non-deterministic.

Under Definition 10, *ceteris paribus*, the foresight exhibited by politicians is more restricted when the environment they face is more electorally competitive. As the incidence of gridlock is increasing with more limited foresight (Corollary 2), we then have a straightforward predicted relationship: *as uncertainty over control of legislative institutions increases, gridlock also increases.* This is among the most important causes of modern-day U.S. gridlock posited by political scientists—seminal work in this literature includes Lee (2016), Binder (1999), among others.

Chapter 2

Endogenous Contribution Cycles*

2.1 Introduction

We consider a class of problems in which a group of agents makes sequential contributions to a project or public good. We focus on *binary* public goods. These are projects that generate no benefits until they have been completed, and then yield a constant benefit flow. While the textbook example of a binary project is the building of a bridge, this class of problems includes a large number of applications, including fundraising for a building, joint team projects with a defined goal, open source software, scientific research to cure a disease, or crowdsourcing.

In some instances, agents have access to an enforcement and commitment technology that allows them to write full contracts regulating their contributions to the public good. More commonly, though, this commitment technology is unavailable, and contributions by different agents occur over time, in a decentralized process.

When this is the case, agents' static free-riding incentives are changed by dynamic considerations. Incentives to contribute are shaped by expectations of contributions by other agents down the line.

*This chapter is joint work with Matias Iaryczower and Santiago Oliveros. I thank Leeat Yariv, Sylvain Chassang, and Silvana Kratseva for their helpful comments. I also thank audiences from the Princeton University political economy and micro-economic theory workshops, as well as the Kansas Workshop on Economic Theory 2021, for their helpful comments. Iaryczower gratefully acknowledges financial support from NSF grant SBE 1757191. Oliveros gratefully acknowledges the support of the Economic and Social Research Council (UK), ES/S01053X/1. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Thus, for instance, a researcher who thinks that her coauthors will not work hard after she makes a contribution to the project might not have incentives to contribute in the first place.

A natural "solution" to this problem is for agents to alternate in making small steps towards completion of the project, dividing the larger project into smaller parts. In fact, as Admati and Perry (1991) point out, this idea goes back to Schelling (1956):

"If each party agrees to send a million dollars to the Red Cross on condition the other does, each may be tempted to cheat if the other contributes first, and each one's anticipation of the other's cheating will inhibit agreement. But if the contribution is divided into consecutive small contributions, each can try the other's good faith for a small price. Furthermore, since each can keep the other on short tether to the finish, no one ever need risk more than one small contribution at a time." (Schelling (1956), p. 45).)

Previous work (Admati and Perry (1991), Compte and Jehiel (2003, 2004)) assumed the alternation takes place, and studied equilibria of the game. In this paper, we do not assume that agents alternate in making contributions, but derive this as a property of equilibrium, in a model with two types of agents, who differ in their valuation of the good. We show that if the project is sufficiently large, the unique equilibrium of the model displays *endogenous contribution cycles*, in which agents of different types alternate making gradual contributions towards the completion of the project. Different from the one-step alternation assumed in the literature, however, in general, type 1 and type 2 agents alternate in the role of active contributors in *stints* of multiple steps. In fact, we show that contribution cycles are generally asymmetric, in the sense that one of the types is an active contributor in stints of multiple steps, while agents of the other type only have single-step stints, except possibly in the final stint before completion.

The alternation of contributions in equilibrium partially solves the commitment problem, but the equilibrium remains inefficient, due to both delay and inaction along the path of play. In particular, we show that not all welfare-enhancing projects are completed, even as players become arbitrarily patient. Moreover, the expected delay is a convex, increasing function of the project size, and goes to infinity as the project size approaches the maximum project size that is completed in equilibrium.

In the model, n agents collaborate to produce a public good, which requires q units of investment. We assume there are no deadlines, or asymmetric information among agents. In each period $t = 1, 2, \ldots$, an agent i selected at random has the opportunity to make a contribution. To distinguish strategic considerations from technological incentives for gradualism, we assume that the cost of contributing is linear. To study in-group and out-group incentives to free ride, we assume there are two types of agents, $m = 1, 2$, and $n^m > 1$ agents of type m. We assume that the probability that any particular individual of type m is selected is $\pi^m = \tilde{\pi}^m / n^m$, where $\tilde{\pi}^1 + \tilde{\pi}^2 = 1$. In line with the binary project assumption, we assume that an agent of type m gets a payoff $\theta^m > 0$ when the project is completed, and zero otherwise. We consider symmetric Markov perfect equilibria (MPE) of the game, in which strategies are a function of the amount outstanding for completion of the good, $s \in (0, q]$.

We begin by analyzing "small" projects. The concept of small is endogenous: we define a project as small if there exists an equilibrium in which the project is completed immediately with positive probability, and denote the size of the largest small project by ϕ_1. We show that small projects have a unique equilibrium. In this equilibrium, agents of a type $m(1) \in \{1, 2\}$ finish the project outright whenever they have an opportunity to move, while agents of type $m' \neq m(1)$ make gradual contributions, or no contributions at all. In particular, there is an *infinite* sequence of cutpoints for type m' agents, $\{s_1(k)\}$, such that whenever $s \in (s_1(k-1), s_1(k)]$, a type 2 agent contributes $e = s_1(k-1) - s$, moving the project to the next cutpoint. Remarkably, in equilibrium $s_1(k) < \phi_1$ for all k, and the sequence of cutpoints converges to a limit $\overline{s}_1 < \phi_1$. Thus there exists a non-empty interval $[\overline{s}_1, \phi_1]$ in which only type $m(1)$ agents contribute on the path of play. In this region, type m' agents display *inaction*. It follows that if $q = \phi_1$, only type $\tilde{m}$ agents contribute on the equilibrium path.

Building on our analysis of small projects, we characterize equilibria for all $q > \phi_1$ recursively. We show that any project $q > \phi_1$ can be seen as a sequence of small projects $\{[\phi_{\tau-1}, \phi_\tau]\}$, each giving agents a payoff equal to their discounted continuation value at completion. From our analysis of small projects, we know that for any $\tau \geq 1$, and any state $s > \phi_\tau$ from which the project

will be completed in equilibrium, the project will eventually reach state $s = \phi_\tau$. At this point, exactly one type of agent, $m(\tau)$, will make a positive contribution, and in fact will complete the project $[\phi_{\tau-1}, \phi_\tau]$, moving the state to $\phi_{\tau-1}$. We call $m(\tau)$ the *active contributor* in step τ. In our main result (Theorem 2), we show that if the project is large enough, type 1 and type 2 agents endogenously alternate in the role of active contibutor, until either the sequence $\{\phi_\tau\}$ converges or q is reached.

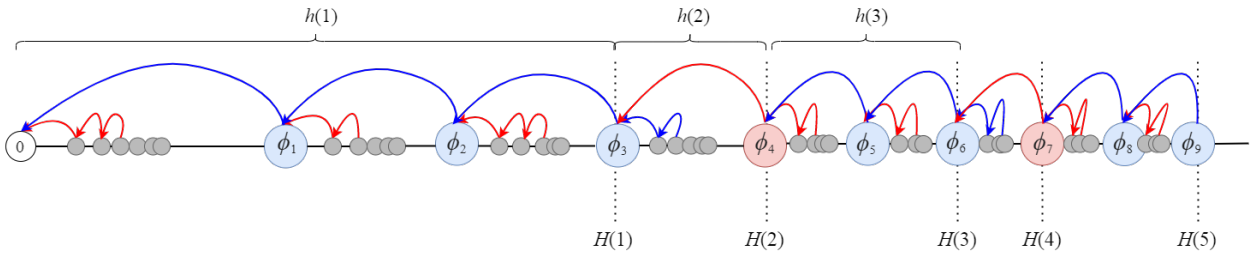


Figure 2.1: Contribution Cycles: Type 1 (blue) and 2 (red) agents alternate contributing in stints of steps towards completion of the project.

Given the structure of equilibrium in small projects, and the fact that the large project can be seen as a sequence of small projects, the main intuition for the existence of contribution cycles is straightforward. First, note that type m agents will be the active contributors at step τ if and only if the discounted value of completing the small project $[\phi_{\tau-1}, \phi_\tau]$ is larger for type m than for type m'. Now, consider two states, s', and $s'' > s'$. Note that agents' equilibrium payoffs at s'' are lower than at s' due to delay, due to impatience, and, *for active contributors*, due to the anticipated cost of moving the project forward. As a result, while the value of completing project $[\phi_\tau, \phi_{\tau+1}]$ is lower than the value of completing project $[\phi_{\tau-1}, \phi_\tau]$ for all agents, the decline is larger for the active contributor in step τ. Since this effect is compounded when type m agents are active contributors for a stint of $h > 1$ steps, the identify of the agent with a higher willingness to pay to complete the small project switches for h large enough.

As the discussion above illustrates, the length of each contribution stint is determined by the number of steps required to change the ranking of the values that type 1 and 2 agents give to completing the next small project. Since agents' preference intensities to complete the project are

arbitrary, the length of the final stint of steps, $h(1)$, is a function of both free riding opportunities and relative preference intensity. At $\phi_{h(1)}$, though, initial differences in preference intensity cancel out with the anticipated cost of the final stint of contributions. As a result, the length of all preceding contribution stints is determined solely by free riding incentives. We show that the group of agents who are active contributors in "long" stints of contribution steps is the group for which in-group free-riding incentives are *larger* (i.e., m such that $n^m > n^{m'}$), and out-group free-riding incentives are *smaller* (i.e., m such that $\tilde{\pi}^m > \tilde{\pi}^{m'}$).

Using our equilibrium characterization, we study the efficiency of equilibrium outcomes. The main result of Compte and Jehiel (2003) is that with linear costs, two players with heterogeneous values, and fixed alternation in play, all welfare-enhancing projects are completed when players are sufficiently patient, and they are completed in at most two steps. We show that a similar result extends to our setting. However, we also show that this result crucially depends on having only one agent of each type (thus eliminating in-group free riding incentives). Whenever this is not the case, some welfare-enhancing projects are not completed, even as players become arbitrarily patient.

2.2 Related Literature

Our paper is related to the literature on dynamic contributions to public goods. The general problem considered by this literature is how free-riding incentives are changed by dynamic considerations.

Two early papers are Fershtman and Nitzan (1991) and Admati and Perry (1991). Fershtman and Nitzan (1991) consider a problem in which the public good delivers continuous benefits, and study the efficiency of public good provision in steady state. Admati and Perry (1991), instead, consider a binary project, in which benefits are attained only when the project is completed, as in our paper. The distinction between continuous and binary public goods is fundamental for dynamic free-riding incentives. As pointed out by Kessing (2007), in continuous public goods, efforts are *strategic substitutes* across time. In this case, contributions lead to a reduction of effort in the future. In binary public goods, instead, efforts are *strategic complements* across time, in that a

higher contribution by one agent induces larger contributions in the future.[1] The latter is a basic feature of our model, as well as of Admati and Perry (1991), Compte and Jehiel (2003, 2004)), Kessing (2007), Giorgiadis (2014), and Bowen et al. (2016). On the other hand, Battaglini et al. (2014) and Harstad (2016), consider models in which the public good gives continuous benefits, as in Fershtman and Nitzan (1991), while Marx and Matthews (2000) consider both cases.

Our paper is most closely related to Admati and Perry (1991) and Compte and Jehiel (2003). Admati and Perry (1991) consider a model in which two homogeneous agents take turns in making contributions to the public good. Admati and Perry show that if the cost of contributing is strictly convex, the game has an essentially unique SPE path, in which agents make gradual contributions to the project.[2] If the contribution cost is linear, the inefficient equilibrium with gradual contributions still exists, although there is also a no-delay equilibrium, in which the project is completed in one step.

Compte and Jehiel (2003) consider a variant of Admati and Perry's model in which agents differ in their valuation of the good. To isolate strategic considerations for gradualism from technological considerations, they focus on the case of linear costs. Compte and Jehiel show that whenever the project is socially desirable, the project is completed as players grow infinitely patient, and in equilibrium, each player makes only one large contribution. Thus, the gradual equilibrium of Admati and Perry appears not to be robust to asymmetries in agents' valuations.[3] Compte and Jehiel conclude that "the sunk character of contributions does not per se provide a satisfactory explanation of the commonly observed contribution pattern in which agents alternate in making small steps towards the completion of the transaction."

[1] This induces inefficient delay due to a time-consistency problem, because contributing more heavily early on and then reducing effort is not credible. In order to prevent others from free riding, each player delays her contributions inefficiently.

[2] This assumes the cost of provision is not "too high"; otherwise the public good is not provided and agents contribute nothing. This threshold is inefficient, in the sense that some socially desirable projects may not be completed.

[3] The key idea behind this result is that with asymmetry in valuations, the player who values the project most is unable to commit not to complete the project whenever the amount left to be contributed is inferior to his valuation. Thus, in equilibrium, the player who values the project most makes the last contribution and the cost of that contribution is equal to his valuation for the project.

Our paper considers a similar environment as that in Compte and Jehiel (2003), but with multiple agents of each type, and no pre-specified order of play. Instead, agents are selected randomly to make a contribution. We show that cycles emerge endogenously in equilibrium if the project is large enough. In this equilibrium, agents make partial contributions to the public good, recovering the gradualism in Admati and Perry (1991). Moreover, we show that our model produces the result of Compte and Jehiel (2003) in which the project is completed in two steps in the special case in which there is exactly one agent of each type.[4] We conclude that gradualism and alternation can appear as a purely strategic phenomenon (without strictly convex costs), provided there are in-group free-riding incentives. The emergence of endogenous contribution cycles is, to the best of our knowledge, new to the literature.

The gradualism present in the equilibrium of our model is captured in the literature through different incentive structures. Compte and Jehiel (2004) consider a variant of the model in Compte and Jehiel (2003) in which each party has the option to terminate the game, and in case of termination, the pay-offs obtained by players depend on the history of offers or concessions made in the bargaining process. They show that history-dependent outside options can lead to gradualism. Kessing (2007) generalizes the gradualism of Admati and Perry (1991) that arises from convex costs to a differential game counterpart, with $n > 1$ homogeneous agents. Differently to most papers in the literature, Marx and Matthews (2000) focus on history-dependent strategies. They show that allowing contributions to be made slowly over time enhances efficiency in some equilibria, even though individual contributions are private information. In these equilibria, players' contributions are supported by the threat of stopping cooperation.[5] Finally, Bonatti and Hörner (2011) consider an "experimentation" model of public good provision. Completing the project requires a breakthrough, which in turn requires effort. Some projects can never be completed, and achieving a breakthrough is the only way to ascertain the project's type. Agents' choice of effort is unobserved by the other agents, which leads them to postpone effort, as in other work reviewed here. However, unlike

[4] In our case, however, the non-active type of agent displays inaction.

[5] A key part of the argument is to show that these strategies can be implemented even when the identity of the deviator is not common knowledge.

these related papers (including our own), the effort expended decreases over time due to growing pessimism.

Furthermore, we consider in our paper how the recognition probabilities of different types and the scale of the project affects equilibrium structure and efficiency. Bowen et al. (2016) consider a problem in which agents choose the scale of the binary project. The analysis for fixed project size builds on the model proposed by Kessing (2007), with two heterogeneous agents. The main contribution of the paper is to analyze how institutions for the collective decision of the scale of the project (unanimity/dictatorship) affect equilibrium outcomes in this setting. Giorgiadis (2014) focuses on institutional design in dynamic contribution games. He studies this problem in a continuous time moral hazard model with simultaneous contributions, in which contributions advance the project stochastically, and effort choices are unobservable.[6]

2.3 Model

We consider a model in which n agents collaborate to produce a public good, which requires q units of investment. There is an infinite horizon, $t = 1, 2, \ldots$. In each period t, an agent i selected at random has the opportunity to invest $e^i(t) \geq 0$ towards the public good. Contributing e units has a cost $C(e) = ce$, for $c > 0$. Letting $e(t)$ denote the contribution made in period t by the agent selected in that period, the amount of effort outstanding for completion of the good at the end of period t is

$$s(t) = \max\{s(t-1) - e(t), 0\}, \quad \text{with } s(0) = q$$

There are two types of agents, $m = 1, 2$, and $n^m > 1$ agents of type $m = 1, 2$.[7] We assume that the probability that any particular individual of type m is selected is $\pi^m = \tilde{\pi}^m / n^m$, where

[6] A key result for his analysis in the underlying contribution game is that members of a larger team work harder than members of a smaller team both individually and on aggregate if and only if the project is sufficiently far from completion.

[7] We postpone the analysis of the case where there is some $j \in \{1, 2\}$ such that $n^j = 1$ to the Remarks below.

$\tilde{\pi}^1 + \tilde{\pi}^2 = 1$. The payoff of an agent i with type $m = 1, 2$ is:

$$U^i = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1}\left[u^m(s(t)) - ce^i(t)\right],$$

where $u^m(s(t)) = 0$ for $s(t) > 0$ and $u^m(0) = \theta^m > 0$.

An *equilibrium* is a symmetric Markov perfect equilibrium (MPE). The state is the amount outstanding for completion of the good, $s \in S = [0, q]$. A pure strategy for player i of type m is a function $\sigma^m(\cdot) : S \to \mathbb{R}_+$, where $\sigma^m(s)$ denotes the investment of an agent of type m in state s when she has the opportunity to contribute.

For any $s \in (0, q]$, the payoff for a type m agent of contributing $e \in [0, s]$ at $s \in S$ when is her time to play is

$$W^m(e, s) = \delta V^m(s - e) + (1-\delta)\left[u^m(s - e) - ce\right], \tag{2.1}$$

where the value of a player i of type m in state s, $V^m(s)$, satisfies *(i)* $V^m(0) = \theta^m$ and *(ii)* for any $s > 0$, letting $\sigma^m(s) = \arg\max_{e \geq 0} W^m(e, s)$,

$$V^m(s) = \sum_{j=1}^{2} \tilde{\pi}^j \left[\delta V^m(s - \sigma^j(s)) + (1-\delta)u^m(s - \sigma^j(s))\right] - \pi^m c(1-\delta)\sigma^m(s) \tag{2.2}$$

We formally define our notion of equilibrium discussed above, as follows.

Definition 11. *A symmetric MPE in pure strategies is a pair* $(\sigma^1(\cdot), \sigma^2(\cdot))$ *such that* (2.1) - (2.2) *hold for all* $s \leq q$

2.4 Results

In this section, we present our results. We begin by analyzing "small" projects in Section 2.4.1. As we discussed in the introduction, the concept of small is endogenous. We define a project as small if there exists an equilibrium in which the project is completed immediately with positive probability, and denote the largest small project by ϕ_1. After characterizing equilibria of small projects, we turn

to our main analysis in Section 2.4.2. We show that any project of size $q > \phi_1$ can be seen as a sequence of small projects. Thus, the equilibrium can be characterized recursively, following the analysis of Section 2.4.1. Using these tools, we show that there exists a unique symmetric MPE, and provide necessary and sufficient conditions for the emergence of endogenous contribution cycles. We close by discussing the efficiency of equilibrium outcomes in Section 2.4.2.

2.4.1 Small Projects

In this section, we analyze equilibrium behavior in small projects. We begin by establishing that in equilibrium, there exists a unique threshold $w(1) > 0$ such that $\sigma^m(s) = s$ for all $m \in M$ if and only if $s \leq w(1)$. Thus, the project is completed without delay if and only if $s \leq w(1)$.

Note first that for s sufficiently small, it is a dominant strategy for any agent to finish the project outright.[8] Thus, there exists a point $w(1)$ such that all agents finish outright whenever $s \leq w(1)$. Now, for each type $m \in M$, let $\hat{\sigma}^m$ denote the function $\hat{\sigma}^m(s) = s$ and define

$$w^m(1) = \max_{s \in S} \{s \in S : W^m(s, s|\hat{\sigma}^m, \hat{\sigma}^{m'}) \geq W^m(e, s|\hat{\sigma}^m, \hat{\sigma}^{m'}) \quad \forall e \neq s\}.$$

The threshold $w^m(1)$ denotes the largest s such that it is a best response for a type m agent to finish the project outright when all other agents do the same. Note that by definition,

$$W^m(s, s|\hat{\sigma}^m, \hat{\sigma}^{m'}) = \theta^m - (1 - \delta)cs.$$

and for $e < s$:

$$W^m(e, s|\hat{\sigma}^m, \hat{\sigma}^{m'}) = \delta \left[\theta^m - \pi^m(1 - \delta)c(s - e)\right] - (1 - \delta)ce,$$

[8] This is intuitive: if the agent does not finish outright, she can at best expect another agent to finish the project tomorrow for sure. Thus, the payoff for waiting is at most $\delta\theta^m$. On the other hand, finishing outright gives a payoff $\theta^m - c(1 - \delta)s$ which is higher than $\delta\theta^m$ whenever $cs < \min\{\theta^1, \theta^2\}$.

which is decreasing in e. Therefore $W^m(s,s|\hat{\sigma}^m,\hat{\sigma}^{m'}) \geq W^m(e,s|\hat{\sigma}^m,\hat{\sigma}^{m'})$ for all $e < s$ if and only if:

$$s \leq \left(\frac{1}{1-\delta\pi^m}\right)\frac{\theta^m}{c} \equiv w^m(1)$$

Let $w(1) \equiv \min\{w^1(1), w^2(1)\}$. Note that by construction, it is an equilibrium for all agents to finish the project outright whenever $s \leq w(1)$. Moreover, since contributions are strategic substitutes, the equilibrium conjecture $(\hat{\sigma}^m, \hat{\sigma}^{m'})$ maximizes free riding incentives for any agent. Thus, this is in fact the unique symmetric MPE for $s \leq w(1)$. Finally, note that if $w^m(1) < w^{m'}(1)$, then by construction, $\sigma^m(s) < s$ whenever $s \in (w(1), w^{m'}(1)]$. Thus for $s > w(1)$ the project is not completed without delay.

Lemma 6. *In a symmetric MPE, $\sigma^m(s) = s$ for all $m \in M$ if and only if $s \leq w(1)$.*

Given the significance of $w^m(1)$ on the construction of equilibrium that follows, for presentation purposes we assume an ordering of agents' willingness to finish the project.

Assumption 4. $w^2(1) < w^1(1)$.

Under assumption 4, as we have seen, $\sigma^2(s) < s$ for any $s > w(1)$. Because contributions are strategic substitutes, this implies that there is $s' > w^1(1)$ such that $\sigma^1(s) = s$ for all $s \leq s'$. Let ϕ_1 denote the largest such s; i.e., the largest state s such that type 1 agents finish the project outright in equilibrium. How would type 2 agents react in states $(w(1), \phi_1]$? We will show that for all states s below some threshold $\overline{s}_1 \leq \phi_1$, in equilibrium type 2 agents use a *cutpoint strategy*, gradually moving the project towards completion when they have an opportunity to play.

Definition 12. *We say that σ^m is a cutpoint strategy in $[a,b] \subseteq S$ if there is a strictly increasing sequence $\{s(k)\}_{k=0}^K$, with K possibly infinite, such that*

1. *$s(0) \equiv a$, and $\lim_{k\to K} s(k) = b$,*
2. *$\sigma^m(s) = s - s(k-1)$, $\forall s \in (s(k-1), s(k)]$ and $\forall k : 1 \leq k \leq K$.*

Let $s_1(0) = 0$. Note that by our previous arguments, in equilibrium type 2 agents follow a cutpoint strategy in $[0, w(1)]$. So let $s_1(1) = w(1)$. Now consider a state $s > w(1)$ close to $w(1)$. We know already that type 2 agents don't finish the project outright whenever $s > w(1)$. In fact it is easy to see that type 2 agents will contribute at most $s - w(1)$: making a contribution that jumps over $w(1)$ without finishing the project doesn't reduce delay, but reduces the maximum benefit of free riding, as any investment above $w(1)$ could be saved with some probability.

On the other hand, investing $e \in (0, s - w(1))$ is also wasteful. *Given that* $\sigma^2(s) \leq s - w(1)$, a type 2 agent is strictly better off by choosing $e = 0$ for s close to $w(1)$. But for all s close to $w(1)$, an agent is strictly better off expending the effort to move the project to $w(1)$ rather than not contributing. This inequality is preserved up to some point $s_1(2) > w(1)$, at which a type 2 agent is indifferent between contributing $e = s - w(1)$ and contributing zero, *given the equilibrium conjecture that* $\sigma^2(s_1(2)) = s_1(2) - w(1)$. It follows that, in equilibrium, type 2 agents follow a cutpoint strategy in $[0, s_1(2)]$. We can then extend this logic recursively for states $s > s_1(2)$, provided at each step k, $s_1(k) < \phi_1$.

The previous argument – which we make precise in the proof of Proposition 8 – shows that in equilibrium, type 2 agents follow a cutpoint strategy for all $s \leq \min\{\phi_1, \lim_{k \to K} s_1(k)\}$. How is this sequence of cutpoints determined?

Note that for all k in the sequence, and all $s \in (s_1(k-1), s_1(k)]$, the value (2.2) for player of type m is

$$V^m(s) = \tilde{\pi}^1 \theta^m + \tilde{\pi}^2 \delta V^m(s_1(k-1)) - \pi^m c(1-\delta)\sigma^m(s) \tag{2.2b}$$

To understand the determination of each cutoff, note that type 2 player's behavior drastically changes around each $s_1(k)$. In fact, when is type 2's turn to contribute at $s = s_1(k) + \epsilon$, it must be that she prefers to contribute in order to reach the cutoff $s_1(k)$ instead of overshooting and contribute to reach any other $s < s_1(k)$. On the other hand, at $s = s_1(k) - \epsilon$, it must be that she prefers to contribute in order to reach the cutoff $s_1(k-1)$ instead of staying in state $s_1(k) - \epsilon$ Taking the limit

as $\epsilon \to 0$ we get

$$\delta V^2(s_1(k)) + c(1-\delta)s_1(k) = \delta V^2(s_1(k-1)) + c(1-\delta)s_1(k-1)$$

which gives

$$\delta V^2(s_1(k)) = \theta^2 - c(1-\delta)s_1(k) \tag{2.3}$$

Combining (2.3) with (2.2b) for $m = 2$ gives

$$s_1(k) = \frac{1}{1-\delta\pi^2}\frac{\theta^2}{c} + \frac{\delta\pi^2}{1-\delta\pi^2}(n^2-1)s_1(k-1) \tag{2.4}$$

Three points are noteworthy. First, note that the equilibrium cutpoints $s_1(k)$ are uniquely determined. Second, note that solving this difference equation allows us to characterize the equilibrium values $V^m(s)$ for both types, for all $s \leq \min\{\phi_1, \lim_{k\to K} s(k)\}$, using equation (2.2). Third, note that

$$s_1(k) - s_1(k-1) = \frac{1}{1-\delta\pi^2}\left(\frac{\theta^2}{c} - (1-\delta\tilde{\pi}^2)s_1(k-1)\right)$$

is decreasing in k. Thus, type 2's contributions become smaller the farther away from completion the project is. This is one form of delay that arises in equilibrium: type 2 agents partially free rides on all types. The further away from completion, the lower the expected value of contributing just a small amount, and the higher the incentives to free ride. To compensate for this and maintain equilibrium, it must be that the cost of contributing decreases, which means reducing the amount of the marginal contribution.

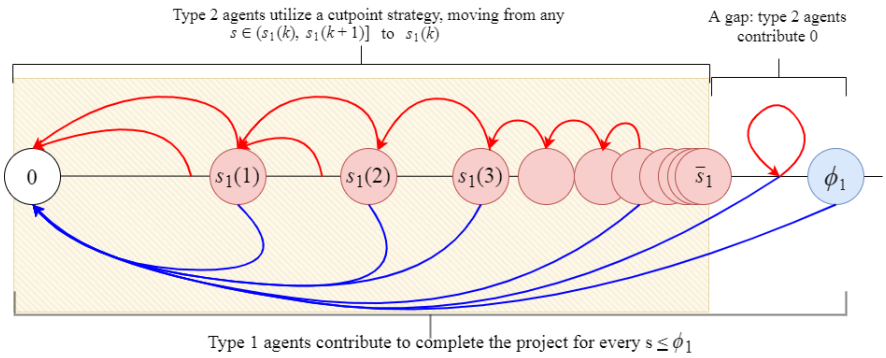


Figure 2.2: Equilibrium in Small Projects, in the generic case $w^1(1) > w^2(1)$.

We show that, remarkably, in equilibrium this process *must come to a halt at a point* $\bar{s}_1$ *strictly below* ϕ_1. Thus, there is a region $[\bar{s}_1, \phi_1]$ within which type 2 agents do not contribute at all, while type 1 agents finish the project outright (see Figure 2.2). This is an extreme form of free riding from type 2 agents, who in this region of the state space fully rely on type 1 agents finishing the project right away. Once this is established, we can then readily obtain a characterization of the threshold ϕ_1, which – since type 2 agents do not contribute at ϕ_1 – does not depend on the sequence of cutpoints $s_1(k)$ for type 2 agents.

Proposition 8. *The unique equilibrium in* $[0, \phi_1]$ *for*

$$\phi_1 = \left(\frac{1}{1-\delta+\delta\pi^1\left(n^1-1\right)}\right)\frac{\theta^1}{c} > \left(\frac{1}{1-\delta\tilde{\pi}^2}\right)\frac{\theta^2}{c} \equiv \bar{s}_1 \tag{2.5}$$

is given by

1. $\sigma^1(s) = s$ *for all* $s \leq \phi_1$

2. σ^2 *is a cutpoint strategy in* $[0, \bar{s}_1] \subset [0, \phi_1]$ *with cutpoints given by*

$$s_1(k) = \frac{1-\left(\delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)\right)^k}{1-\delta\tilde{\pi}^2}\frac{\theta^2}{c} < \bar{s}_1 \quad \forall k \geq 1, \tag{2.6}$$

3. $\sigma^2(s) = 0$ *for all* $s \in [\bar{s}_1, \phi_1]$.

and the values for player type m at all $s \in (s_1(k-1), s_1(k)] \leq \overline{s}_1$ *are given by (2.2b) and at all* $s \in (\overline{s}_1, \phi_1]$ *are given by*

$$V^m(s) = \frac{\tilde{\pi}^1}{1 - \delta\tilde{\pi}^2}\left(\theta^m - \frac{c}{n^m}(1-\delta)\sigma^m(s)\right). \tag{2.7}$$

To understand why $\overline{s} < \phi_1$ consider $\hat{k} \in \{1, 2, \ldots, ...\}$ as the maximal finite index of the sequence $\{s_1(k)\}_{k=1}^{\hat{k}}$ at which player type 2 contributes to reach the next $s_1(\hat{k}-1)$. Let $\tilde{s}(1, \hat{k})$ be the state at which type 1 is indifferent between finishing outright and not contributing, given that agent 2 has exactly $\hat{k}$ cutpoints below $\tilde{s}(1, \hat{k})$. Furthermore, assume that there is an equilibrium in which $s_1(\hat{k}) < \tilde{s}(1, \hat{k}) \leq s_1(\hat{k}+1)$, a necessary condition for the cutoff strategy to be composed of finite cutoffs. We show by construction that this is not possible.

Given that $w^1(1) > w^2(1) = w(1)$, there exists at least one cutpoint of player 2 strictly below $w^1(1)$ as we explained above. Consider then $\hat{k} = 1$ and note that, in fact, $\tilde{s}(1,1)$ must also be greater than $w^1(1)$. Since $w^1(1)$ is the maximal point at which any type 1 completes the project *given that* agents of type 2 also do so at $w^1(1)$, and this is not the case in equilibrium at $w^1(1)$, the previous argument for the existence of $\tilde{s}$ implies that $\tilde{s}(1,1) > w^1(1)$.

Note that once $\hat{k} < \infty$ is fixed, and the assumption about equilibrium $s_1(\hat{k}) < \tilde{s}(1, \hat{k}) \leq s_1(\hat{k}+1)$ in place, it is straightforward to calculate the value functions for this equilibrium candidate. We then use the calculated value functions to determine $\tilde{s}(1, \hat{k})$ explicitly. As $\hat{k}$ increases, the incentives for type 1 to complete the project are compounded, as type 2 agents take progressively smaller steps, pushing $\tilde{s}(1, \hat{k})$ farther away. We show that in fact $s_1(\hat{k}+1) < \tilde{s}(1, \hat{k})$ for all $\hat{k}$, contradicting our assumption about equilibrium. It follows that in equilibrium we must have an infinite sequence of cutpoints strictly below ϕ_1.

There are three key take-away points from Proposition 8. First, small projects have a unique equilibrium. Second, in equilibrium, agents of type $\tilde{m} = \arg\min_m\{w^m(1)\}$ finish the project outright whenever they have an opportunity to move, while agents of type $m' \neq \tilde{m}$ make gradual contributions, or no contributions at all. Third, the size of the gradual contributions by type m'

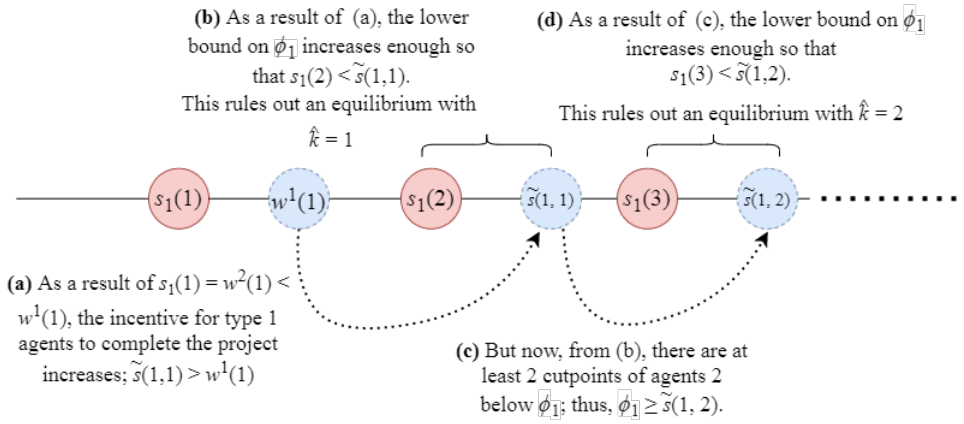


Figure 2.3: Eliminating the possibility of Finitely Many Cutpoints Below ϕ_1.

agents vanishes as we move farther away from completion, and eventually these agents do not contribute at all. Thus in $[\overline{s}_1, \phi_1]$, only type $\tilde{m}$ agents contribute on the path of play. Given this, the values at the "beginning" of the small project are given by

$$V^1(\phi_1) = \left(\frac{(1 - 1/n^1)\tilde{\pi}^1}{1 - \delta + (1 - 1/n_1)\delta\tilde{\pi}^1} \right) \theta^1 \quad \text{and} \quad V^2(\phi_1) = \left(\frac{1 - \tilde{\pi}^2}{1 - \delta\tilde{\pi}^2} \right) \theta^2 \tag{2.8}$$

At the core of these results is the tradeoff between delay and free-riding for each agent. To understand how this tradeoff is resolved in equilibrium, it is useful to decompose free-riding incentives into two distinct components: *in-group* effects and *out-group* effects. Agents in our model can free ride on peers, and on the other group. The incentive to free-ride on peers is the in-group effect, and is parametrized by n^m. The incentive to free-ride on members of the other group is the out-group effect, and is parameterized by $\tilde{\pi}^m$.

Consider first in-group effects. In principle, in-group effects can have both direct and indirect equilibrium effects. Thus, a change in n^1 can affect the equilibrium strategy of type 2 agents and vice-versa. Provided the change in parameters does not affect the ranking $w^1(1) > w^2(1)$, however, in-group effects only affect the behavior of agents of that type in small projects. In fact, note that n^1 affects ϕ_1, but not the sequence of cutpoints $\{s_1(k)\}$, or its limit $\overline{s}_1$. Similarly, for any k finite, the cutpoint $s_1(k)$ is a function of n^2, but neither $\overline{s}_1$ nor ϕ_1 are a function of n^2.

With this noted, the effect of group size on equilibrium outcomes in small projects is straightforward. Reducing n^1 (reducing in-group free riding incentives) increases the maximum size of the project that type 1 agents are willing to complete outright, ϕ_1. As the number of type 1 agents decreases, there is a higher chance that any single type 1 agent would have to contribute in the next period, were she to shirk in the current period. Since at ϕ_1 a type 1 agent must be indifferent in equilibrium between shirking and finishing outright, the lower benefit of shirking must have the effect of increasing the amount of effort required to finish the project. As we can see from (2.25), both the direct and the indirect equilibrium effects of reducing n^1 have a negative impact on type 1 agents' value at the cutpoint, lowering $V^1(\phi_1)$. For exactly the same reason, reducing n^2 increases the size of each increment $s_1(k) - s_1(k-1)$ along the sequence of cutpoints for type 2 agents (see equation (2.4)). However, because ϕ_1 is not affected by n^2, both $V^1(\phi_1)$ and $V^2(\phi_1)$ are unchanged.

Next, we consider out-group effects. Possibly counterintuitively at first, increasing $\tilde{\pi}_1$ (nominally reducing out-group free riding incentives) leads to a lower ϕ_1, thus reducing the size of the largest project that type 1 agents are willing to complete outright, ϕ_1. To understand why this is the case, note that since type 1 agents are the only ones contributing at ϕ_1, the beneficial effect of free-riding to *out-group* agents is quashed. Indeed, in this case, there is no action from type 2 agents to free-ride *on*. Quite to the contrary, each period a type 2 agent is selected, progress on the project remains stalled. Thus, as $\tilde{\pi}^1$ increases, *delay* decreases, lowering the cost of shirking. Hence, ϕ_1 must decrease in order to restore type 1's indifference. As a result, increasing $\tilde{\pi}_1$ increases $V^1(\phi_1)$ *and* $V^2(\phi_1)$.

On the other hand, since type 1 agents complete the project outright for all $s \leq \phi_1$, increasing $\tilde{\pi}_2$ has the effect of increasing all cutpoints $s_1(k)$, as expected. Because the prospect of free riding on type 1 agents is less likely, shirking becomes less attractive. Since at any cutpoint $s_1(k)$ type 2 agents must be indifferent between shirking or moving to the next cutpoint, $s_1(k-1)$, the cost of contributing must increase, increasing $s_1(k)$. Since this holds for all previous cutpoints in the sequence, the effect of increasing $\tilde{\pi}^2$ compounds over time shifting $s_1(k)$ more the higher k is. Thus,

in equilibrium,

$$\frac{ds_1(k)}{d\tilde{\pi}^2} > \frac{ds_1(k-1)}{d\tilde{\pi}^2} > 0 \quad \forall k \geq 1.$$

Before moving on to efficiency considerations, we consider the special case of single-agent teams.

Remark 2 (**Single-Agent Teams**). *So far, we have assumed that $n^m > 1$ for all $m \in M$. The basic structure of equilibria remains essentially unchanged if there is only one agent of any given type. However, equilibrium outcomes have some noteworthy features. The case $n^1 = 1 < n^2$ is unremarkable, with the exception that letting $n^1 = 1$ maximizes the set of projects that are completed immediately with positive probability. Consider next $n^2 = 1 < n^1$. Note that by* (2.6)*, $s_1(k) = w(1) = \overline{s}_1$ for all $k \in \mathbb{N}$. In this case, the single type 2 agent takes one step whenever $s \leq \overline{s}_1$, and remains inactive in $[\overline{s}_1, \phi_1]$. Since $\overline{s}_1$ and ϕ_1 are not a function of n^2, making $n^2 = 1$ has an unambiguous effect of reducing delay in small projects. Not surprisingly, when $n^2 = 1 = n^1$ we have a combination of both cases: more projects are completed immediately with positive probability, and there is less delay in states close to completion.*

Efficiency

In Proposition 8 we showed that for any small project with $q \in (w(1), \phi_1]$, in equilibrium there is delay with positive probability. Efficiency, on the other hand, requires that the project be completed if and only if total benefits exceed total costs,

$$\sum_m n^m \theta^m \geq c(1-\delta)q \Leftrightarrow q \leq \frac{n^1\theta^1 + n^2\theta^2}{c(1-\delta)},$$

and that whenever the project is completed, it is completed outright. Since $w(1) = \left(\frac{1}{1-\delta\pi^2}\right)\frac{\theta^2}{c}$ it follows immediately that there is a nonempty set of projects for which the equilibrium is inefficient. Given our equilibrium characterization, in fact, it is straightforward to compute the expected time

for completion of the project from any state $s \leq \phi_1$, which is given by

$$\mathcal{E}(s) = \begin{cases} \frac{1-(\tilde{\pi}^2)^k}{\tilde{\pi}^1} & \text{if } s \in (s_1(k-1), s_1(k)], \text{ for } k = 1, \dots \\ \frac{1}{\tilde{\pi}^1} & \text{if } s \in [\overline{s}_1, \phi_1] \end{cases}. \tag{2.9}$$

Using (2.6) to substitute the cutpoints in terms of primitives, it follows that for any project $q < \overline{s}_1$,

$$\mathcal{E}(q) = \frac{1-(\tilde{\pi}^2)^{\psi(q)}}{\tilde{\pi}^1} \quad \text{where} \quad \psi(q) \equiv \left\lceil \frac{\log\left(1 - \frac{cq}{\theta^2}(1-\delta\tilde{\pi}^2)\right)}{\log\left(\delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)\right)} \right\rceil \tag{2.10}$$

Note that $\psi(q)$, and therefore $\mathcal{E}(q)$, is weakly increasing in the total cost of the project cq, and weakly increasing in the intensity of preferences of type 2 agents for the good, θ^2. Moreover, as we discussed before, increasing n^2 reduces the size of each increment $s_1(k) - s_1(k-1)$ along the sequence of cutpoints for type 2 agents, and does not affect $\overline{s}_1$ or ϕ_1. As a result, expected delay for $q < \overline{s}_1$ is also weakly increasing in the number of type 2 agents. Interestingly, expression (2.10) shows that there is an upper bound on expected delay for any $q \leq \phi_1$, given by $1/\tilde{\pi}^1$. Thus, even when a change of parameters leads to a large increase in the k such that $q \in (s_1(k-1), s_1(k)]$, this does not lead to a proportional increase in the expected time for completion.

2.4.2 Large Projects

Having characterized the equilibrium of the game for "small projects"; i.e., $s \leq \phi_1$, we move to studying the equilibrium of the contribution game for any possible project length q. Our first result justifies the attention we paid to understanding small projects. We show that any project $q > \phi_1$ can be seen as a sequence of small projects. Thus, the equilibrium can be characterized recursively, following the analysis of Section 2.4.1.

Theorem 1. *There exists an almost everywhere unique symmetric MPE in pure strategies. The equilibrium is characterized by an increasing sequence* $\{\phi_\tau\}_{\tau=0}^{\overline{\tau}}$*, with* $\phi_0 = 0$ *and* $\overline{\tau}$ *possibly infinite, such that:*

(i) For any $\tau \geq 1$, there is a unique $m_\tau \in \{1,2\}$ such that

(a) $\sigma^{m_\tau}(s) = s - \phi_{\tau-1}$ for all $s \in (\phi_{\tau-1}, \phi_\tau]$, and

(b) σ^{-m_τ} is a cutpoint strategy in $[\phi_{\tau-1}, \overline{s}_\tau]$ for some $\overline{s}_\tau < \phi_\tau$, and $\sigma^{-m_\tau} = 0$ in $(\overline{s}_\tau, \phi_\tau]$

(ii) If $\{\phi_\tau\}$ converges to a limit $\tilde{\phi} < q$, $\sigma^m(s) = 0 \; \forall s \in [\tilde{\phi}, q]$ and all m.

(iii) For $\tau > 1$,

$$\phi_\tau - \phi_{\tau-1} = \left(\frac{1}{1 - \delta + \delta \pi^{m_\tau}(n^{m_\tau} - 1)} \right) \frac{\delta V^{m_\tau}(\phi_{\tau-1})}{c}, \tag{2.11}$$

The basic intuition for this result relies on two key facts. First, note that at any $s > \phi_1$, any type $m = 1, 2$, is not willing to contribute more than the needed amount to reach ϕ_1, i.e. $\sigma^m(s) \leq s - \phi_1$. It follows that no agent moves the project beyond ϕ_1 at any $s > \phi_1$ (see Lemma 12 in the Appendix). This implies that for the project to be completed from $s > \phi_1$ the project at some point reaches ϕ_1 and has to transit through ϕ_1 on the equilibrium path. It follows then that the game can be truncated at

$$\phi_1 = \left(1 + \frac{\delta(1 - \tilde{\pi}^1)}{1 - \delta\pi^1 - \delta(1 - \tilde{\pi}^1)} \right) w^1(1)$$

by considering ϕ_1 as the terminal state, which pays out $\delta V^m(\phi_1)$ when ϕ_1 is attained.

Second, recall that the equilibrium characterization in Proposition 8 relied solely on the Assumption 4, $w^2(1) < w^1(1)$. In a similar fashion, when considering the truncated game at ϕ_1, we define

$$w^m(2) \equiv \frac{1}{1 - \delta\pi^m} \frac{\delta V^m(\phi_1)}{c} = w^m(1) \frac{\delta V^m(\phi_1)}{\theta^m}$$

Proposition 8 gives that $m_2 = \{m = 1, 2 : w^m(2) > w^{-m}(2)\}$ is the active contributor in this new step, where behavior in this second step is described by replacing type 1 for type m_2, $w^m(1)$ for $w^m(2)$, and θ^m for $V^m(\phi_1)$ (see Figure 2.4).

Note that generically, $w^1(\tau) \neq w^2(\tau)$ for any $\tau \geq 1$, so we can continue to characterize the equilibrium step by step for the full game, by recursively truncating the game at each state ϕ_τ that

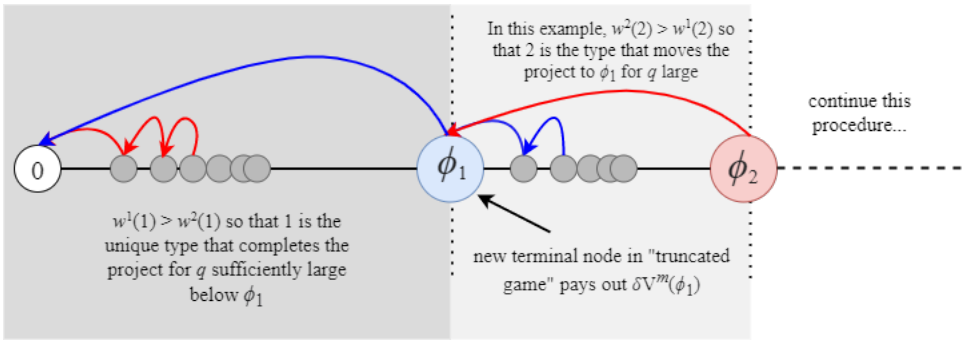


Figure 2.4: Successively truncating the game at the maximal point at which either type is willing to contribute to the previous terminal node, and applying our equilibrium characterization for small projects, we obtain the equilibrium structure of Theorem 1 recursively.

determines the upper bound of step τ for a given $\phi_{\tau-1}$.[9] Continuing this process, we continue to obtain the same structure recursively. Proceeding in this fashion, parts *(i)* and *(iii)* of the theorem follow immediately from our analysis in Section 2.4.1. For *(ii)*, note that if $\phi_\tau \to \tilde{\phi} < q$, then for each $m \in \{1, 2\}$,

$$w^m(\tau) := \frac{\delta V^m(\phi_{\tau-1})}{c(1-\delta\pi^m)} \to 0,$$

for else, $\phi_\tau \geq \phi_{\tau-1} + \min\{w^1(\tau), w^2(\tau)\}$, which would contradict convergence of $\{\phi_\tau\}$. Hence, $V^m(\phi_\tau) \to 0$, so that $V^m(\tilde{\phi}) = 0$. It follows that at any $s > \tilde{\phi}$, contributing nothing is strictly prefer by both types to contributing any positive amount.

Having established the fact that large projects can be seen as a sequence of small projects, we move on to consider the possible emergence of contribution cycles. From Theorem 1, there is generically a unique contributor at each cutpoint ϕ_τ. We call $m(\tau)$ the *active contributor* in step τ, and let $-m(\tau) = \{m \in M : m \neq m(\tau)\}$. From (2.25), and using $\delta V^m(\phi_1)$ as the terminal payoffs in the truncated game, for all $\tau > 1$ we have:

$$V^{m_\tau}(\phi_\tau) = \alpha^{m_\tau} V^{m_\tau}(\phi_{\tau-1}) \quad \text{and} \quad V^{-m_\tau}(\phi_\tau) = \beta^{-m_\tau} V^{-m_\tau}(\phi_{\tau-1}) \tag{2.12}$$

[9] Generically projects are finished in finite steps or contributions never start. In the former, generically $w^m(\tau) \neq w^{m'}(\tau)$, while in the latter this equality cannot be ruled out by standard measure theoretic arguments. Nevertheless, in any equilibrium with $\overline{\tau} \to \infty$, the case in point, contributions never start generically. In that sense, there could be multiple equilibria that sustain all equilibrium in which projects are never finished, and hence, $V^m(q) = 0$.

where

$$\alpha^m \equiv \frac{\delta\pi^m(n^m-1)}{\delta\pi^m(n^m-1)+(1-\delta)} \quad \text{and} \quad \beta^m \equiv \frac{\delta(1-\tilde{\pi}^m)}{1-\delta\tilde{\pi}^m}.$$

Clearly, both $\alpha^m < 1$ and $\beta^m < 1$, so agents' values decrease as we move farther away from completion. Moreover, from (2.12),

$$\frac{V^{m_\tau}(\phi_\tau)}{V^{-m_\tau}(\phi_\tau)} = \Delta^{m_\tau}\frac{V^{m_\tau}(\phi_{\tau-1})}{V^{-m_\tau}(\phi_{\tau-1})} \quad \text{where} \quad \Delta^m \equiv \frac{\alpha^m}{\beta^{-m}} < 1, \tag{2.13}$$

reflecting the fact that the value of the active contributor decreases faster than that of the non-active contributor. Note that the equilibrium payoff of non-active contributors only decreases with τ due to impatience and delay, which occurs when one of them is selected to contribute. Thus their value decreases with τ faster the smaller is the probability that a contributor is selected to contribute, $\tilde{\pi}^{m_\tau}$, and the lower is the discount factor δ. On the other hand, the contributor's value decreases with τ due to impatience, delay, *and* due to the cost of moving the project forward. Since this expected cost is smaller the larger the within-group externalities, their value decreases slower the larger n^{m_τ} is.

Since the value of the active contributors decreases with τ faster than the value of the non-active contributors, whenever the same type plays the role of active contributor in successive steps, initial differences in willingness to pay at any given state ϕ_τ wash out the farther away from ϕ_τ the process is. This opens the possibility that the identity of the active contributor may change after a sufficiently large number of steps.

For any $\tau \geq 1$, define $\Omega^m(\phi_\tau) \equiv w^m(\phi_\tau)/w^{-m}(\phi_\tau)$. Note that $m(\tau) = m$ if and only if $\Omega^m(\phi_\tau) > 1$. Since $w^m(\tau) = \frac{\delta V^m(\phi_{\tau-1})}{c(1-\delta\pi^m)}$ is linear in $V^m(\phi_{\tau-1})$, we have

$$\Omega^{m_\tau}(\phi_{\tau+1}) = \frac{V^{m_\tau}(\phi_\tau)}{V^{-m_\tau}(\phi_\tau)} \times \chi(m_\tau) \quad \text{and} \quad \Omega^{m_\tau}(\phi_\tau) = \frac{V^{m_\tau}(\phi_{\tau-1})}{V^{-m_\tau}(\phi_{\tau-1})} \times \chi(m_\tau),$$

for a constant $\chi(m_\tau)$. Thus, the growth rate of $\Omega^{m_\tau}(\phi_\tau)$ is equal to the growth rate of $V^{m_\tau}(\phi_{\tau-1})/V^{-m_\tau}(\phi_{\tau-1})$, and then by (2.13), we have:

$$\Omega^{m_\tau}(\phi_{\tau+1}) = \Delta^{m_\tau}\Omega^{m_\tau}(\phi_\tau). \tag{2.14}$$

The previous expression allows us to determine the identity of the active contributor, m_τ, recursively. Note that by construction we must have $\Omega^{m_\tau}(\phi_\tau) > 1$. Given this, $m_{\tau+1} = m_\tau$ if and only if $\Omega^{m_\tau}(\phi_{\tau+1}) > 1$, or equivalently $\Delta^{m_\tau}\Omega^{m_\tau}(\phi_\tau) > 1$. We use this result to study the emergence of contribution cycles. We begin by providing a necessary and sufficient condition for agents of different types to alternate in the role of active contributors.

Recall that assumption (4) implies that $m_1 = 1$. Suppose type 1 agents remain in the role of active contributors for a stint of $h(1) \geq 1$ consecutive steps. Noting that $\Omega^1(\phi_\tau) = (\Delta^1)^{\tau-1}\Omega^1(\phi_1)$ for all $\tau \leq h(1)+1$, there is a switch in the identify of the active contributor at $h(1)+1$ if and only if

$$\Omega^1(\phi_{h(1)+1}) < 1 < \Omega^1(\phi_{h(1)}) \quad \Leftrightarrow \quad (\Delta^1)^{h(1)} < \Omega^2(\phi_1) < (\Delta^1)^{h(1)-1}$$

Since $\Delta^1 < 1$, and by assumption $\Omega^2(\phi_1) < 1$, there exists a unique $h(1) \in \mathbb{Z}_+$ that satisfies this inequality, which is given by

$$0 < h(1) - \frac{\log w^2(1) - \log w^1(1)}{\log(\Delta^1)} < 1$$

This means that the final stint in which type 1 agents are the active contributors is preceded by a stint of type 2 agents in the role of active contributors if and only if q is large enough.

Proposition 9. *Type 1 agents are the only active contributors on the equilibrium path; i.e., $m_\tau = 1$ for all $\tau : \phi_\tau \leq q$, if and only if*

$$q < \left(\frac{1}{1 - \delta\tilde{\pi}^2 - \delta\tilde{\pi}^1/n^1}\right)\frac{1-(\alpha^1)^{h(1)}}{(1-\alpha^1)}\frac{\theta^1}{c} = \phi_{h(1)},$$

Proposition 9 says that when $q < \phi_{h(1)}$, only type 1 agents are active contributors on the equilibrium path. While this *does not* necessarily imply that only type 1 agents contribute on the path of play, it does rule out contribution cycles. To see this, it is useful to focus on equilibrium outcomes. Suppose first that $q = \phi_\tau$ for some $\tau \leq h(1)$. Then only type 1 agents contribute on the path of play, finishing the project in τ steps. The project transitions from ϕ_τ to $\phi_{\tau-1}$, from $\phi_{\tau-1}$ to

$\phi_{\tau-2}$, etc, with possible delay at each point, whenever type 2 agents have an opportunity to move. The exact same outcome occurs if $q \in [\overline{s}_\tau, \phi_\tau]$, as type 2 agents do not contribute in this interval, and type 1 agents move the project to $\phi_{\tau-1}$. Finally, suppose $\phi_{\tau-1} < q < \overline{s}_\tau$. In this case, type 1 agents move the project to $\phi_{\tau-1}$ outright, as before, but type 2 agents make small contributions, moving the project gradually closer to $\phi_{\tau-1}$. Irrespective of the particular sequence of agents selected to play, eventually $\phi_{\tau-1}$ is reached, and the process continues as in the previous cases, with only type 1 agents making contributions on the path of play.

Provided $q > \phi_{h(1)}$, the final stint of h_1 steps in which type 1 agents act as active contributors is preceded by a stint of steps in which type 2 agents are active contributors. How does this process evolve for q large? Suppose type 2 agents remain in the role of active contributors for a stint of $h(2) \geq 1$ consecutive steps. Note that for all $\tau : h(1) + 1 \leq \tau \leq h(1) + h(2) + 1$:

$$\Omega^1(\phi_\tau) = \frac{1}{(\Delta^2)^{\tau-h(1)-1}}\Omega^1(\phi_{h(1)+1}) = \frac{(\Delta^1)^{h(1)}}{(\Delta^2)^{\tau-h(1)-1}}\Omega^1(\phi_1).$$

Thus, there is a switch in the identify of the active contributor at $h(1) + h(2) + 1$ if and only if

$$\Omega^1(\phi_{h(1)+h(2)}) < 1 < \Omega^1(\phi_{h(1)+h(2)+1}) \quad \Leftrightarrow \quad \frac{(\Delta^1)^{h(1)}}{(\Delta^2)^{h(2)-1}} < \Omega^2(\phi_1) < \frac{(\Delta^1)^{h(1)}}{(\Delta^2)^{h(2)}}$$

Since $\Delta^1 < \Omega^2(\phi_1)/(\Delta^1)^{h(1)-1} < 1$ by definition of $h(1)$, there is a $h(2) \geq 1$ that satisfies this inequality. This alternation between type 1 and type 2 agents in the role of active contributors continues until the associated sequence of contributions $\phi_\tau - \phi_{\tau-1}$ in each step τ converges to zero, or the cumulative contributions reach q (whichever happens earlier). This gives the main result of the paper:

Theorem 2 (Contribution Cycles). *For any integer $k \geq 1$, define recursively the total number of steps up to and including stint k as $H(k) \equiv H(k-1) + h(k)$, letting $H(0) = 0$. Let $\iota(k) = 1$ if k*

is odd, $\iota(k) = 2$ *if* k *is even, and* $\Omega^1(\phi_0) \equiv \frac{1-\delta\pi^2}{1-\delta\pi^1}\frac{\theta^1}{\theta^2}$*. Then for all* $k \geq 1$*:*

$$m_\tau = \begin{cases} 1 & \text{if } H(2k-2) < \tau \leq H(2k-1) \\ 2 & \text{if } H(2k-1) < \tau \leq H(2k), \end{cases}$$

where the number of steps in stint k*,* $h(k)$*, is given recursively by:*

$$\Omega^{\iota(k)}(\phi_{H(k)}) = [\Delta^{\iota(k)}]^{h(k)}\Omega^{\iota(k)}(\phi_{H(k-1)}) \qquad \text{and} \qquad \Delta^{\iota(k)} < \Omega^{\iota(k)}(\phi_{H(k)}) < 1. \tag{2.15}$$

Theorem 2 shows that in the generically unique equilibrium, type 1 and type 2 agents endogenously alternate actively contributing to the public good, until either the sequence $\{\phi_\tau\}$ converges or q is reached. This defining feature of equilibrium endogenizes the alternation in player contributions that was heretofore exogenously imposed in the related literature, following Admati and Perry (1991). To our knowledge, our paper is the first to demonstrate the emergence of *endogenous contribution cycles* in equilibrium.

In general, each contribution stint k can have multiple steps. However, in the special case in which $\Delta_1 = \Delta_2 = \Delta$, the ratio of values changes at the same rate when type 1 agents and type 2 agents are active contributors. Therefore, for any $k \geq 1$ odd we have:

$$\Omega^1(\phi_{H(k)}) = [\Delta]^{\sum_{r=3}^{k}(h(r)-h(r-1))}\Omega^1(\phi_{H(1)}).$$

As a result, asymmetries in the length of the contribution stints by each type lead to imbalances in the ratio of values which are inconsistent with equilibrium. In fact, in this case we must have that types alternate after a single step after the final stint of contribution steps by type 1 agents. To see this formally, note that from Theorem 2, at any k odd for equilibrium we need

$$\Delta^1 < \Omega^1(\phi_{H(k)}) < 1 \quad \text{and} \quad \Delta^2\Omega^1(\phi_{H(k)}) < [\Delta^2]^{h(k+1)} < \Omega^1(\phi_{H(k)}).$$

If $\Delta^1 \geq \Delta^2$, we get that $(\Delta^2)^{h(k+1)} > (\Delta^2)^2$, and hence $h(k+1) = 1$. On the other hand, if $\Delta^1 \leq \Delta^2$, we get that $h(k+2) = 1$.

Now, note that

$$\Delta^m = \frac{(n^m - 1)(1 - \delta\tilde{\pi}^{-m})}{n^m(1 - \delta\tilde{\pi}^{-m}) - \delta\tilde{\pi}^m}$$

Therefore, if groups are symmetric, so that $n_m = n/2$ and $\tilde{\pi}_m = 1/2$ for all $m \in M$, then $\Delta_1 = \Delta_2 = \Delta$. We then obtain the following result:

Corollary 4. *If both groups have equal in-group and out-group free-riding incentives, then* $h(k) = 1$ *for all* $k > 1$.

When $\Delta_1 \neq \Delta_2$, the rate of change of $\Omega^m(\phi_\tau)$ varies with the identity of the active contributor, and one-step cycles may not be sufficient to maintain a balance in the relative willingness to pay. In order to restore equilibrium, then, the length of the stints by each type must be asymmetric. In our next result, we characterize the asymmetry in contribution cycles.

Proposition 10. *If* $\Delta^{m'} < \Delta^m$*, contribution cycles alternate between single-step stints by type* m' *agents and stints of either* x^m *or* $x^m + 1$ *steps by type* m *agents, where*

$$x^m \equiv \left\lfloor \frac{\log[\Delta^{m'}]}{\log[\Delta^m]} \right\rfloor$$

Two points are worth emphasizing. First, note that the length of each stint $k > 1$ only depends on free-riding incentives, and *not on relative preference intensity*. This contrasts with the very last stint of steps by player 1 ($k = 1$). In fact, note that from Theorem 2, $h(1)$ is given by

$$\Delta^1 < [\Delta^1]^{h(1)} \frac{1 - \delta\pi^2}{1 - \delta\pi^1} \frac{\theta^1}{\theta^2} < 1.$$

The reason for this is fundamentally endogenous: all asymmetry in willingness to pay due to differences in preference intensity washes out in the last contribution stint by type 1 agents, and is thus solely reflected in $\phi_{h(1)}$. After this point, the determination of the identity of the active contributor solely depends on free-riding incentives, as captured by n^m and $\tilde{\pi}^m$.

Second, note that if $n^1 = n^2 = n/2$, then $\Delta^1 > \Delta^2$ if and only if $\tilde{\pi}^1 > 1/2$. Similarly, if $\tilde{\pi}^1 = \tilde{\pi}^2 = 1/2$, then $\Delta^1 > \Delta^2$ if and only if $n^1 > n^2$. The first case balances out in-group free riding incentives. Proposition 10 says that the group of agents who are active contributors in "long" stints of contribution steps is the group for which out-group free-riding incentives are *smaller*. While this result might appear counterintuitive at first glance, the logic is familiar to us from the analysis of small projects. What determines the length of a contribution stint is how fast the ratio of values $V^{m_\tau}(\phi_{\tau-1})/V^{-m_\tau}(\phi_{\tau-1})$ decreases towards balance when type m is the active contributor. As we showed in Section 2.4.1, the reason for the "reversal" of the sign of the out-group effect on type m agents is that at the cutpoint $\phi(\tau)$, type $-m$ agents don't make positive contributions. Thus, shifting recognition probability to the opposite group effectively does not increase free-riding opportunities for type m agents, and increases delay. As a result, increasing $\tilde{\pi}^{m_\tau}$ leads to a lower depreciation of the value of the active contributors. The second case balances out-group effects across groups. Proposition 10 says that the group of agents who are active contributors in "long" stints of contribution steps is precisely the group for which in-group free-riding incentives are larger. The reason for this is straightforward. A larger group size means a smaller expected dent in the value of the active contributors in each step, because other agents of the same type can also foot the bill. Thus, it takes more steps to achieve balance in the ratios of values, $V^{m_\tau}(\phi_{\tau-1})/V^{-m_\tau}(\phi_{\tau-1})$.

The previous results fully characterize equilibria of the game. In Section 2.4.2, we analyze the efficiency of equilibrium outcomes. Before doing so, we consider equilibria of the game in the special case in which one or both groups are composed of a single member, so there are no in-group free riding incentives.

Single-Agent Teams. Suppose $n^1 = n^2 = 1$. As discussed in Section 2.4.1 the equilibrium for small projects in this case is characterized by cutpoints

$$s_1(1) = \frac{\theta_2/c}{1 - \delta\tilde{\pi}_2} \quad \text{and} \quad \phi_1 = \frac{\theta_1/c}{1-\delta}$$

such that

$$\sigma^1(s) = s \quad \forall s \leq \phi_1 \quad \text{and} \quad \sigma^2(s) = \begin{cases} s & \forall s \leq s_1(1) \\ 0 & \forall s \in (s_1(1), \phi_1] \end{cases}$$

Since at the cutpoint ϕ_1 the single type 1 agent is indifferent between completing the project and not contributing, we must have $V^1(\phi_1) = 0$. This implies that for all $s > \phi_1$, only type 2 contributes. It is easy to see then that $\sigma^2(s) = s - \phi_1$ for all $s \in (\phi_1, \phi_2]$ with

$$\phi_2 = \phi_1 + \frac{\delta\tilde{\pi}_1}{1 - \delta\tilde{\pi}_2}\frac{\theta_2/c}{1-\delta}$$

and $V^2(\phi_2) = 0$. It follows that the project is completed if and only if

$$q \leq \frac{1}{1-\delta}\left[\frac{\theta_1}{c} + \frac{\delta(1-\tilde{\pi}_2)\theta_2}{c(1-\delta\tilde{\pi}_2)}\right], \tag{2.16}$$

and if completed, is completed in at most two steps.

This result is analogous to the main theorem of Compte and Jehiel (2003) (see Proposition 1 in that paper). We can think of the above result as extending their finding to the case where recognition among two heterogeneous agents is stochastic instead of sequentially determined. Importantly, Theorem 3 shows that this result depends crucially on there being a single agent of each type, and thus on the non-existence of in-group free-riding incentives.

Next, suppose $n^2 > 1$ and $n^1 = 1$. As in the previous case, here $\phi_1 = (\theta_1/c)/(1-\delta)$ and $V^1(\phi_1) = 0$, and the single type 1 agent is unwilling to contribute for any $s > \phi_1$. The game truncated at final node ϕ_1 now only involves type 2 agents, who are fully homogeneous, as in Admati and Perry (1991). The equilibrium of the truncated game is characterized by the sequence

$$\phi_\tau = \phi_1 + \frac{\delta\tilde{\pi}^1}{1-\delta\tilde{\pi}^1}\frac{\Delta^2}{(n^2-1)}\left[\frac{1-(\Delta^2)^\tau}{1-\Delta^2} - 1\right]\overline{s}_1 \qquad \text{for } \tau > 1$$

with associated values

$$V^2(\phi_\tau) = (\Delta^2)^{\tau-1}\frac{\tilde{\pi}_1}{1-\delta\tilde{\pi}_2}\theta_2,$$

recovering the structure of the equilibrium in Admati and Perry (1991).

Efficiency

We now analyze the efficiency of equilibrium outcomes. We consider two distinct questions related to efficiency: (i) whether all welfare-improving projects are completed, and (ii) what is the delay incurred in finishing efficient projects.

Our normalization of utility by factor $1-\delta$ affects both the project's infinite-lasting flow of payoffs and the one-stage cost of each paid contribution. This normalization simplifies the expressions of the value function as θ^m represents both a flow payoff and the discounted utility of the project. For any $\delta < 1$ this double role of θ does not affect the arguments above and, in particular, does not affect the efficiency considerations. However, as $\delta \to 1$, any finite contribution in a single period vanishes relative to the discounted utility of the infinite-lasting project. Hence, when discounting disappears, efficiency considerations are more subtle. To simplify the expressions in this section, we consider a project that, once completed, lasts for T finite periods and delivers θ^m in each period. The utility of such a finished project to agent i of type $m = 1, 2$ is

$$U^i = (1-\delta)\sum_{t=1}^{T} \delta^{t-1} u^m(s(t)) = (1-\delta)\hat{\theta}^m,$$

and we note that

$$\lim_{\delta\to 1} \frac{U^i}{1-\delta} = \lim_{\delta\to 1} \hat{\theta}^m = \lim_{\delta\to 1} \frac{1-\delta^T}{1-\delta}\theta^m = T\theta^m.$$

Hence, in general, the payoff of an agent i with type $m = 1, 2$ is:

$$U^i = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} \left[u^m(s(t)) - ce^i(t)\right],$$

where $u^m(s(t)) = 0$ for $s(t) > 0$ and $u^m(0) = \hat{\theta}^m > 0$.

Implementable Projects. By expression (2.11) in Theorem 1, for any $\tau > 1$, if type m is the unique contributor at τ,

$$\phi_\tau - \phi_{\tau-1} = \left(\frac{1}{1 - \delta\tilde{\pi}^{-m} - \delta\pi^m} \right) \frac{\delta V^m(\phi_{\tau-1})}{c}.$$

where

$$V^m(\phi_{\tau-1}) = \begin{cases} \alpha^m V^m(\phi_{\tau-2}) & \text{if } m(\tau-1) = m \\ \beta^m V^m(\phi_{\tau-2}) & \text{if } m(\tau-1) \neq m \end{cases}$$

Since $\alpha_m < \beta_m < 1$, the value $V_m(\phi_\tau)$ is decreasing in τ, and goes to zero as $\tau \to \infty$. It follows that the incremental contribution $\phi_\tau - \phi_{\tau-1}$ is decreasing in τ and goes to zero as $\tau \to \infty$. This means that the sequence $\{\phi_\tau\}$ converges to a point $\tilde{\phi}$. The threshold $\tilde{\phi}$ gives the size of the smallest project that will not be completed in equilibrium.

Now, note that using (2.11) and the derived values for each agent at the end of any given contribution stint, we can calculate $\phi_{H(k)}$ for any $k \geq 1$. In particular, note that the values at the end of any given stint k only depend on the number of times each agent contributes up to the k^{th} stint. For instance, for $k \geq 1$ odd, we have

$$V^1(\phi_{H(k)}) = [\alpha^1]^{h(k)}[\beta^1]^{h(k-1)} V^1(\phi_{H(k-2)})$$

and

$$V^2(\phi_{H(k)}) = [\beta^2]^{h(k)}[\alpha^2]^{h(k-1)} V^2(\phi_{H(k-2)})$$

Proposition 10 then allows us to obtain an upper bound for the limit $\tilde{\phi}$. We present the exact characterization of this bound in Lemma 13 in the Appendix. Using this bound, we address the question of whether all projects which would be efficient to pursue – i.e., projects such that $cq < (n^1\hat{\theta}^1 + n^2\hat{\theta}^2)$ – are in fact completed. We call these projects *efficient* for short. We show that in equilibrium not all efficient projects are completed, even in the limit as $\delta \to 1$.

Theorem 3. *For any $\delta < 1$, $c\tilde{\phi} < \hat{\theta}^1 + \hat{\theta}^2$. Hence, (i) for any fixed $\delta < 1$, some efficient projects are not completed. Moreover, (ii) since $n_m > 1$ for any $m \in \{1, 2\}$, then even in the limit as $\delta \to 1$, some efficient projects are not completed.*

The qualification that $n_m > 1$ for any $m \in \{1, 2\}$ is important. As we discussed in the previous section, when $n_1 = n_2 = 1$ the project is completed if and only if (2.16) holds, and if completed, it is completed in at most two steps. It follows that

$$c\tilde{\phi}_{n^1=n^2=1} = \hat{\theta}_1 + \frac{1 - \tilde{\pi}_2}{1 - \delta\tilde{\pi}_2}\delta\hat{\theta}^2$$

which implies that in the limit as $\delta \to 1$, all efficient projects are completed.

Delay. As we have shown before, the incremental contribution $\phi_\tau - \phi_{\tau-1}$ in (2.11), is decreasing in τ and goes to zero as the sequence $\{\phi_\tau\}$ approaches the limit $\tilde{\phi}$. It follows that the number of steps required to finish a project of size q, $H(\hat{k}(q))$, where $\hat{k}(q) \equiv \min\{k : \phi_{H(k)} \geq q\}$, (i) is an increasing and convex function of q, and (ii) $H(\hat{k}(q)) \to \infty$ as $q \to \tilde{\phi}$. Because the number of steps in each contribution stint is fixed, this means that the number of contribution stints required to complete a project of size q, $\hat{k}(q)$, is also an increasing and convex function of q, such that $\hat{k}(q) \to \infty$ as $q \to \tilde{\phi}$.

In our next result, we use the structure of equilibrium strategies characterized in the previous section to compute the expected delay $\mathcal{E}(q)$ associated with completion of a project of size q. In particular, we first show that the expected completion time of a project that requires $\hat{k}(q)$ contribution stints, $\{h(1), \ldots, h(\hat{k}(q))\}$, i.e. $\phi_{\hat{k}(q)-1} < q \leq \phi_{\hat{k}(q)}$, is given by

$$\mathcal{E}\left(\{h(k)\}_{k=1}^{\hat{k}(q)}\right) = \sum_{k=1}^{\hat{k}(q)} (\tilde{\pi}^{\iota(k)})^{h(k)-1}\left[h(k) + \frac{\tilde{\pi}^{\iota(k+1)}}{\tilde{\pi}^{\iota(k)}}\right],$$

where $\iota(k) = 1$ if k is odd and $\iota(k) = 2$ if k is even. By Proposition 10, moreover, we know that if $\Delta^{m'} < \Delta^m$, contribution cycles alternate between stints of either x^m or $x^m + 1$ steps by type

m agents, and single-step stints by type m' agents, while if $\Delta^{m'} = \Delta^m$, both types alternate in single-step stints. This gives tight bounds on expected delay for any project size q.

Proposition 11. *Suppose $q = \phi_\ell > \phi_{h(1)}$, and let $\mathcal{E}_1 \equiv (\tilde{\pi}^1)^{h(1)-1}\left[h(1) + \frac{\tilde{\pi}^2}{\tilde{\pi}^1}\right]$. If $\Delta^m > \Delta^{m'}$, then the expected delay $\mathcal{E}(q)$ associated with completion of a project of size q is:*

$$\mathcal{E}(q) = \mathcal{E}^1 + \left(\frac{\hat{k}(q)}{2}\right)\left[(\tilde{\pi}^m)^{x-1}\left[x + \frac{\tilde{\pi}^{m'}}{\tilde{\pi}^m}\right] + \frac{1}{\tilde{\pi}^{m'}}\right]$$

for some $x \in [x^m, x^m + 1]$, and if $\Delta^m = \Delta^{m'}$, then:

$$\mathcal{E}(q) = \mathcal{E}^1 + \left(\frac{\hat{k}(q)}{2}\right)\left[\frac{1}{\tilde{\pi}^1} + \frac{1}{\tilde{\pi}^2}\right].$$

In particular, $\mathcal{E}(\cdot)$ is an increasing and convex function, and $\lim_{q\to\tilde{\phi}} \mathcal{E}(q) = \infty$.

Proposition 11 quantifies delay from any member of the sequence $\{\phi_\ell\}$. Therefore, together with the characterization for delay in small projects given by (2.10), we have a full characterization of delay from arbitrary q. Proposition 11 connects the two components of efficiency discussed in this paper. The first, inherently tied with the structure of equilibrium we characterized, is gradualism and the associated inefficient delay, The second, from theorem 3, is the result that some efficient projects are not completed, even as $\delta \to 1$. Proposition 11 shows that the expected delay increases non-linearly with the size of the project, and goes to infinity as the size of the project approaches the smallest project not completed in equilibrium, $\tilde{\phi}$. Thus, the expected delay is a smooth measure of both types of inefficiency.

2.5 Conclusion

In this paper, we study sequential contributions to binary public goods, in a decentralized environment in which commitment to a contribution schedule is not feasible. We consider a model in which agents are selected at random to make a contribution to the project each period until the project is

completed. We assume there are no deadlines or asymmetric information among agents. We show that if the project is sufficiently large, the unique equilibrium of the model displays endogenous contribution cycles, in which agents of different types alternate making gradual contributions towards the completion of the project.

Contribution cycles emerge in our model as a key characteristic of a gradual process of contributions featuring inaction by some agents on the path of play. In particular, we show that any project of size q large enough is equivalent to completion of a series of small projects $\{[\phi_{\tau-1}, \phi_\tau]\}$, such that completion of each small project gives agents a payoff equal to their discounted continuation value at completion. Each small project is completed outright by an agent of a type $m(\tau)$, the active contributor, while agents of type $-m(\tau)$ remain inactive, inducing further delay.

Due to the lack of commitment, the active contributor in each small project is determined in a recursive process akin to backward induction, by the relative value each type gives to completing the small project. Because contributing is costly, taking the role of the active contributor in project $[\phi_{\tau-1}, \phi_\tau]$ reduces the value that type $m(\tau)$ agents give to completing $[\phi_\tau, \phi_{\tau+1}]$ relative to completing the project $[\phi_{\tau-1}, \phi_\tau]$ by a *larger* amount than for type $-m(\tau)$ agents. Since this effect compounds as the same type of agents are responsible for larger stints of contributions, there must eventually be a switch in the identity of the active contributor, inducing contribution cycles.

As we discussed in the Introduction, our model covers a number of applications, including fundraising for a building, joint team projects with a defined goal, open source software, scientific research to cure a disease, or crowdsourcing. Many of the insights of our model can also be extended to related problems, such as gradualism in multilateral bargaining (following the insight of Compte and Jehiel (2004)).[10] In this context, our results provide an explanation for gradualism in mulitalteral bargaining that does not depend on history-dependent outside options.

[10]Suppose there are x dollars to distribute from reaching agreement. Each agent of type m is a stakeholder with a claim θ^m, where the total claims $y = \sum n^1\theta^1 + n^2\theta^2$ exceed the value of agreement, so that $y - x = q > 0$. Each agent i can agree to some cut to her claims $e^i \geq 0$. If $\sum e^i = q$, agents get to distribute the proceeds, otherwise they get zero. Note that this model is isomorphic to the one we considered in the paper. Extending this model to the case in which contributions are paid only upon reaching a deal would be natural in some contexts, as in bankruptcy restructuring proceedings.

In the paper, we focused on the case of two types of agents for simplicity. We conjecture that the main insights of the model extend to the case in which there are $L > 2$ types of agents, each with a different valuation for the good. Suppose in this setting, too, that a unique type contributes at the largest small project, ϕ_1 (this is intuitive, but would need to be proven). We conjecture that the equilibrium would have the following feature. Exactly one type of agent, say 1, is the active contributor in a final stint of steps. This eventually brings the value of completion of a small project for type 1 agents "close to" the value for type 2 agents, yet still above that for other types, $\ell = 3, \ldots, L$. Proceeding backwards, type 1 and 2 agents alternate taking steps, until their value is close to that for type 3 agents, and still above that for all other types. At that point, type 1, 2, and 3 agents alternate making contributions, until their value approaches the value for the next type. Continuing in this fashion, eventually all types alternate making contributions until the process converges, or q is reached.

Our paper shows that endogenous contribution cycles provide a partial solution to commitment problems in dynamic public good provision games. The "solution," however, is second best. The equilibrium is still inefficient, due to both delay and inaction along the path of play. Moreover, the expected delay increases more than proportionally with the size of the project, and goes to infinity as the project size approaches the maximum project size that is completed in equilibrium. This leaves a large space for institutions to improve the efficiency of collective action. The optimal design of these mechanisms in this context is an attractive avenue for future research.

Bibliography

Admati, A. R. and Perry, M. (1991). Joint projects without commitment. *The Review of Economic Studies*, 58(2):259–276.

Battaglini, M., Nunnari, S., and Palfrey, T. R. (2014). Dynamic free riding with irreversible investments. *American Economic Review*, 104(9):2858–2871.

Bonatti, A. and Hörner, J. (2011). Collaborating. *American Economic Review*, 101(2):632–63.

Bowen, T. R., Georgiadis, G., and Lambert, N. S. (2016). Collective choice in dynamic public good provision. Working Paper.

Compte, O. and Jehiel, P. (2003). Voluntary contributions to a joint project with asymmetric agents. *Journal of Economic Theory*, 112(2):334 – 342.

Compte, O. and Jehiel, P. (2004). Gradualism in bargaining and contribution games. *The Review of Economic Studies*, 71(4):975–1000.

Fershtman, C. and Nitzan, S. (1991). Dynamic voluntary provision of public goods. *European Economic Review*, 35:1057–1067.

Giorgiadis, G. (2014). Projects and team dynamics. *The Review of Economic Studies*, 82(1):187–218.

Harstad, B. (2016). The dynamics of climate agreements. *Journal of the European Economic Association*, 14(3):719–752.

Kessing, S. G. (2007). Strategic complementarity in the dynamic private provision of a discrete public good. *Journal of Public Economic Theory*, 9(4):699–710.

Marx, L. M. and Matthews, S. M. (2000). Dynamic voluntary contribution to a public project. *The Review of Economic Studies*, 67:327–358.

Schelling, T. C. (1956). An essay on bargaining. *The American Economic Review*, 46(3):281–306.

2.A Proofs for Chapter 2

2.A.1 Proof of Lemma 6

(a) First we show that there exists a $b > 0$ such that $\sigma^m(s) = s$ for all m for all $s \leq b$. Note that for $s \leq \min\{\theta^1, \theta^2\}/c$,

$$W^m(s, s) := \theta^m - (1 - \delta)cs \geq \delta\theta^m, \ \forall m,$$

where the right-hand side is an upper bound on $W^m(e, s)$ for any $e < s$, since it is the value from zero contribution and the project being completed by another agent in the subsequent period. Letting $b = \min\{\theta^1, \theta^2\}/c$ completes the step.

(b) We now determine the candidates $w^m(1)$. It is clear that for $s > \min^m\{\frac{\theta^m}{c(1-\delta)}\}$ we can never have $\sigma^m(s) = s$ for all m, for this would give some agents a negative payoff, making it profitable to deviate and contribute zero at s. Thus, there exists a finite

$$w(1) \equiv \sup\{b : \sigma^m(s) = s \ \forall s \leq b, \forall m \in M\}$$

Note that for $e < s$:

$$W^m(e, s) = \delta V^m(s - e) - (1 - \delta)ce.$$

Take any $s \leq w(1)$. From (2.2), and since $\sigma^m(s) = s$ for all m for $s < w(1)$, for any $e < s$ we have[11]

$$V^m(s - e) = \theta^m - \pi^m c(1 - \delta)(s - e)$$

Thus for $e < s$:

$$W^m(e, s) = \delta\theta^m - c(1 - \delta)\left(e + \delta\pi^m(s - e)\right)$$

Note that for any $0 < e < s$:

$$W^m(0, s) - W^m(e, s) = c(1 - \delta)e\left(1 - \delta\pi^m\right) > 0$$

Thus, when all agents finish the project on their turn, an agent is better off contributing zero than contributing $e \in (0, s)$. Moreover, $W^m(s, s) \geq W^m(0, s)$ if and only if

$$s \leq \frac{\theta^m}{c\left(1 - \delta\pi^m\right)} \equiv w^m(1),$$

with equality uniquely at $s = w^m(1)$. Thus $\sigma^m(s) = s$ for $m = 1, 2$ if $s < w(1) \equiv \min^m w^m(1)$ and only if $s \leq w(1)$. In particular, we see it is always an equilibrium that $\sigma^m(w(1)) = w(1)$ for $m = 1, 2$.

(c) We consider, moreover, conditions for the uniqueness of equilibrium behavior at $w(1)$. Note that if $w^m(1) > w(1)$, then agents of type m have a strict incentive to contribute at $w(1)$ and so we need only consider those types m for which $w^m(1) = w(1)$. From previous work, we know that $\sigma^m(w(1)) \in \{0, w(1)\}$. Consider therefore an equilibrium in which $\sigma^m(w(1)) = 0$. In such an

[11] Note that it is possible that $\sigma^m(w(1)) \neq w(1)$. However, in this case because $w(1)$ is the supremum of the states at which m contributes, she is indifferent between contributing and not. Hence, when $s = w(1)$ and $e = 0$, $V^m(s - e)$ will have the same value as that given in the above proof.

equilibrium,

$$V(w(1)) = \frac{\tilde{\pi}^{m'}\theta^m}{1-\delta\tilde{\pi}^m}. \tag{2.17}$$

Consider now a deviation of a player of type m to contributing $w(1)$,

$$\begin{aligned} W^m(w(1), w(1)) &= \theta^m - c(1-\delta)w(1) = \delta\theta^m \frac{1-\pi^m}{1-\delta\pi^m} \\ &= \delta V(w(1)) \frac{\frac{1-\pi^m}{1-\delta\pi^m}}{\frac{1-\tilde{\pi}^m}{1-\delta\tilde{\pi}^m}} > \delta V(w(1)) \end{aligned} \tag{2.18}$$

Thus, there is a profitable deviation for the type m agents from the conjectured equilibrium and so we have that equilibrium behavior at $w(1)$ is unique and entails completion of the project.

(c) It remains to show that if $w^m(1) > w^{m'}(1)$, then $\sigma^m(s) = s$ for all $s \in (w^{m'}(1), w^m(1)]$. So let consider first $s \in (w^{m'}(1), w^m(1))$. As before, $W^m(s,s) = \theta^m - c(1-\delta)s$. However, since type m' agents don't finish the project outright, for any $e \in [0, s)$ we have

$$W^m(e,s) < \delta\theta^m - c(1-\delta)\left[e + \delta\pi^m(s-e)\right] \leq \delta\left(\theta^m - c(1-\delta)\pi^m s\right)$$

Thus $W^m(s,s) > W^m(e,s)$ for all $e \in [0,s)$ if and only if $s < w^m(1)$. Now consider the situation when $s = w^m(1)$. Using similar arguments we have that $W^m(w^m(1), w^m(1)) > W^m(e, w^m(1))$ for all $e \in (0, w^m(1))$, so we only need to consider equilibria with $\sigma^m(w^m(1)) \in \{0 w^m(1)\}$. The definition of $w^m(1)$ gives that when $\sigma^m(w^m(1)) = w^m(1)$ is indeed an equilibrium. For uniqueness assume that $\sigma^m(w^m(1)) = 0$. It follows then that

$$V^m(w^m(1)) = \frac{\tilde{\pi}^{m'}\delta}{1-\tilde{\pi}^m\delta} V^m\left(w^m(1) - \sigma^{m'}(w^m(1))\right)$$

Since no type finishes the project outright, we must have that

$$V^m\left(w^m(1) - \sigma^{m'}(w^m(1))\right) \leq \theta^m - \pi^m c(1-\delta)\left(w^m(1) - \sigma^{m'}(w^m(1))\right)$$

Therefore, there is a profitable deviation for type m if $\delta V^m(w^m(1)) < \theta^m - c(1-\delta)w^m(1)$. It is sufficient if

$$\begin{aligned} \frac{\tilde{\pi}^{m'}\delta}{1-\tilde{\pi}^m\delta}\left(\theta^m - \pi^m c(1-\delta)\left(w^m(1) - \sigma^{m'}(w^m(1))\right)\right) &< \theta^m - c(1-\delta)w^m(1) \\ \frac{1-\tilde{\pi}^m}{n^m - \tilde{\pi}^m}\sigma^{m'}(w^m(1)) &< w^m(1) \end{aligned}$$

which is true. □

2.A.2 Proof of Proposition 8

We prove this in a series of steps

Lemma 7. *Suppose $s_1(2)$ is such that $\sigma^1(s) = s$ for all $s \leq s_1(2)$. Then:*

1. *For any* $s \in (s_1(1), s_1(2)]$, $\sigma^2(s) = s - w(1)$.

2. *For any* $s > s_1(2)$, $\sigma^2(s) \leq s - s_1(2)$.

Proof of Lemma 7. From (2.2), it follows that at any cutpoint $s_1(k)$ in the sequence, a type 2 agent's ex ante value is:

$$V^2(s_1(k)) = \tilde{\pi}^2 \delta V^2(s_1(k-1)) + (1 - \tilde{\pi}^2)\theta^2 - \pi_2 c(1-\delta)(s_1(k) - s_1(k-1)) \tag{2.19}$$

and must verify (2.3)

$$\delta V^2(s_1(k)) = \theta^2 - c(1-\delta)s_1(k)$$

Consider the problem of a type 2 agent, i. We know already that for $s > w(1)$, $\sigma^2(s) < s$.

(i) Note that since in equilibrium all agents finish the project outright for $s \leq w(1)$, if agent i contributes $e = s - w(1)$, she gets a payoff

$$\begin{aligned} W^2(s - w(1), s) &= \delta V^2(w(1)) - (1-\delta)c(s - w(1)), \\ &= \delta \left[\theta^2 - \pi^2(1-\delta)cw(1)\right] - (1-\delta)c(s - w(1)). \end{aligned}$$

(ii) Contributing $e \in (s - w(1), s)$ is clearly dominated, since by moving to $w(1)$ the project will be finished at the same time, with lower expected costs.

(iii) Suppose i contributes $e \in [0, s - w(1))$, say $e = s - w(1) - \varepsilon$, with $0 < \varepsilon \leq s - w(1)$. This gives her a payoff

$$W^2(e, s) = \delta V^2(w(1) + \varepsilon) - (1-\delta)c(s - w(1) - \varepsilon)$$

Note that for any $s \in (w(1), w_1(1)]$,

$$V^2(w(1) + \varepsilon) = \tilde{\pi}^2 \delta V^2(w(1) + \varepsilon - \sigma^2(w(1) + \varepsilon)) + (1 - \tilde{\pi}^2)\theta^2 - \pi^2 c(1-\delta)\sigma^2(w(1) + \varepsilon)$$

Since $e \in (s - w(1), s)$ is clearly dominated by $e = s - w(1)$, the best possible outcome is that $\sigma^2(w(1) + \varepsilon) = \varepsilon$. Therefore

$$\begin{aligned} V^2(w(1) + \varepsilon) &\leq \tilde{\pi}^2 \delta V^2(w(1)) + (1 - \tilde{\pi}^2)\theta^2 - \pi^2 c(1-\delta)\varepsilon \\ &= \tilde{\pi}^2 \delta \theta^2 + (1 - \tilde{\pi}^2)\theta^2 - \pi^2 c \left[\tilde{\pi}^2 \delta(1-\delta)w(1) + (1-\delta)\varepsilon\right] \end{aligned}$$

and thus, for $e = s - w(1) - \varepsilon$ ($0 < \varepsilon \leq s - w(1)$):

$$W^2(e, s) \leq \delta \tilde{\pi}^2 \delta \theta^2 + \delta(1 - \tilde{\pi}^2)\theta^2 - \delta(1-\delta)\pi^2 c \left[\tilde{\pi}^2 \delta w(1) + \varepsilon\right] - (1-\delta)c(s - w(1) - \varepsilon)$$

Thus contributing $s - w(1)$ beats contributing $s - w(1) - \varepsilon$ ($0 < \varepsilon \leq s - w(1)$) if

$$\delta \tilde{\pi}^2 \theta^2 \geq \delta \pi^2 c w(1) \left(1 - \tilde{\pi}^2 \delta\right) + c\varepsilon \left(1 - \delta\pi^2\right)$$

Note the RHS is increasing in ε. Thus, this inequality holds for all $\varepsilon \in (0, s - w(1)]$ if it holds for $\varepsilon = s - w(1)$ or, rearranging, if

$$s \leq \frac{1}{c}\left(\frac{n^2}{n^2 - \delta\tilde{\pi}^2}\right)\theta^2 + \delta\tilde{\pi}^2\left(\frac{n^2 - 1}{n^2 - \delta\tilde{\pi}^2}\right)w(1) \equiv s_1(2)$$

It follows that for any $s \in (w(1), s_1(2)]$, $\sigma^2(s) = s - w(1)$. Given this, in turn, the bound we used for $V^2(w(1)+\varepsilon)$ above is actually tight. It follows that $\sigma^2(s) < s - w(1)$ for all $s > s_1(2)$. Now, for any such s, moving the project to any $s \in (s_1(2), w_1)$ is clearly dominated by moving the project to $s_1(2)$, since the project will be finished at the same time, with lower expected costs. Thus $\sigma^2(s) \leq s - s_1(2)$ for all $s > s_1(2)$.

□

Lemma 8. *Suppose there is $b \in \mathbb{R}_+$ such that type 2 agents follow a cutpoint strategy in $[0, b]$ and $\sigma^1(s) = s$ at all $s \leq b$. Then*

1. *The sequence of cutpoints is described by (2.6):*

$$s_1(k) = \frac{1 - \left(\delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)\right)^k}{1-\delta\tilde{\pi}^2}\frac{\theta^2}{c}$$

2. *The values for type 1 are given by*

$$V^1(s_1(1)) = \theta^1 - \pi_1 n^1 c(1-\delta)s_1(1), \tag{2.20}$$

and for $k > 1$,

$$\begin{aligned} V^1(s_1(k)) = &\left(\frac{1-(\tilde{\pi}^2\delta)^k}{1-\tilde{\pi}^2\delta}\right)\left[\tilde{\pi}^1\theta^1\right] + (\tilde{\pi}^2\delta)^k\theta^1 \\ &- \pi^1\left(\frac{1-\delta}{1-\delta\tilde{\pi}^2}\right)\left[\left(\frac{1-(\delta\tilde{\pi}^2)^k}{1-\delta\tilde{\pi}^2}\right) - \left(\delta\tilde{\pi}^2\right)^k\left(\frac{n^2-\delta\tilde{\pi}^2}{1-\delta\tilde{\pi}^2}\right)\left[1-\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)^k\right]\right]\theta^2 \end{aligned} \tag{2.21}$$

Proof of Lemma 8. Using (2.3) for $s_1(k)$ and $s_1(k-1)$ in (2.19) gives

$$\begin{aligned} s_1(k) &= \frac{1}{c}\left(\frac{n^2}{n^2-\delta\tilde{\pi}^2}\right)\theta^2 + \delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)s_1(k-1) \\ &= w_2(1) + \delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)s_1(k-1) \end{aligned} \tag{2.22}$$

Solving the difference equation, we have:

$$s_1(k) = \left[\sum_{\ell=0}^{k-1}\left(\delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)\right)^\ell\right]\left[\frac{1}{c}\left(\frac{n^2}{n^2-\delta\tilde{\pi}^2}\right)\theta^2\right]$$

and using that $\sum_{\ell=0}^{k-1} x^\ell = \frac{1-x^k}{1-x}$ we get (2.6).

The values of type 1 agents for $k = 1$ follows trivially from the fact that both players contribute to complete the project at $s_1(1)$. Consider $k > 1$. Note that the value for type 1 at $s_1(k)$ is

$$V^1(s_1(k)) = \tilde{\pi}^1\theta^1 + \tilde{\pi}^2\delta V^1(s_1(k-1)) - \pi^1 c(1-\delta)s_1(k) \tag{2.23}$$

and

$$V^1(s_1(k-1)) = \tilde{\pi}^1\theta^1 + \tilde{\pi}^2\delta V^1(s_1(k-2)) - \pi^1 c(1-\delta)s_1(k-1)$$

So substituting,

$$V^1(s_1(k)) = (1+\delta\tilde{\pi}^2)\tilde{\pi}^1\theta^1 - \pi^1 c(1-\delta)\left[\tilde{\pi}^2\delta s_1(k-1) + s_1(k)\right] + (\tilde{\pi}^2\delta)^2 V^1(s_1(k-2)).$$

and

$$V^1(s_1(k-2)) = \tilde{\pi}^1\theta^1 + \tilde{\pi}^2\delta V^1(s_1(k-3)) - \pi^1 c(1-\delta)s_1(k-2)$$

so substituting,

$$V^1(s_1(k)) = \left(\sum_{\ell=0}^{2}(\tilde{\pi}^2\delta)^\ell\right)\left[\tilde{\pi}^1\theta^1\right] - \pi^1 c(1-\delta)\left[\sum_{\ell=0}^{2}(\tilde{\pi}^2\delta)^\ell s_1(k-\ell)\right] + (\tilde{\pi}^2\delta)^3 V^1(s_1(k-3))$$

Doing this recursively until the final term is a multiple of $V^1(s_1(1))$,

$$V^1(s_1(k)) = \left(\frac{1-(\tilde{\pi}^2\delta)^{k-1}}{1-\tilde{\pi}^2\delta}\right)\tilde{\pi}^1\theta^1 + (\tilde{\pi}^2\delta)^{k-1}\theta^1 - \pi^1 c(1-\delta)\left[\sum_{\ell=0}^{k-1}(\tilde{\pi}^2\delta)^\ell s_1(k-\ell)\right]$$

and substituting $s_1(k-\ell)$, we get

$$\begin{aligned} V^1(s_1(k)) = &\left(\frac{1-(\tilde{\pi}^2\delta)^{k-1}}{1-\tilde{\pi}^2\delta}\right)\tilde{\pi}^1\theta^1 + (\tilde{\pi}^2\delta)^{k-1}\theta^1 \\ &- \pi^1(1-\delta)\theta^2\left[\sum_{\ell=0}^{k-1}(\tilde{\pi}^2\delta)^\ell\left(\frac{1-\left(\delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)\right)^{k-\ell}}{1-\delta\tilde{\pi}^2}\right)\right] \end{aligned}$$

simplifying further, we have the expression in the statement of the lemma.

□

Lemma 9. *In any symmetric MPE, $\sigma^2(s)$ is a cutpoint strategy in $[0,\overline{s}_1]$ for $\overline{s}_1 \equiv \sup_{k\geq 1} s_1(k)$ if $\sigma^1(s) = s$ for all $s \leq \overline{s}_1$.*

Proof of Lemma 9. First note that by Lemma 6 and Lemma 7, the statement is true up to $k = 2$, with $s_1(0) \equiv 0$, $s_1(1) \equiv w(1)$, and $s_1(2)$ as defined in lemma 7.

Next, suppose the statement is true up to k, and consider $s \in (s_1(k), s_1(k+1)]$.

(i) First, we show that for any $s \in (s_1(k), s_1(k+1)]$, $\sigma^2(s) \leq s - s_1(k)$. Specifically, we show that for any $s \in (s_1(k), s_1(k+1)]$, if $\sigma^2(s) > s - s_1(k)$, a type 2 agent would prefer to deviate and contribute zero. Note that since moving to a non-cutpoint is dominated by moving to a cutpoint, it is enough to consider $\sigma^2(s) = s - s_1(k-\ell)$ for all $\ell : 1 \leq \ell < k$.

In the proposed equilibrium, a type 2 agent that gets to contribute would get a payoff

$$W^2(s - s_1(k-\ell), s) = \delta V^2(s_1(k-\ell)) - (1-\delta)c(s - s_1(k-\ell))$$

If instead she deviated to contributing zero, she would get

$$W^2(0,s) = \delta V^2(s)$$

Since by assumption $\sigma^2(s) = s - s_1(k-\ell)$ and we know $\sigma^1(s') = s'$ for all $s' \leq \overline{s}_1$, at s we have

$$V^2(s) = \tilde{\pi}^2 \delta V^2(s_1(k-\ell)) + (1-\tilde{\pi}^2)\theta^2 - \pi^2 c(1-\delta)[s - s_1(k-\ell)]$$

Therefore

$$W^2(0,s) = \tilde{\pi}^2 \delta^2 V^2(s_1(k-\ell)) + (1-\tilde{\pi}^2)\delta\theta^2 - \pi^2 c\delta(1-\delta)[s - s_1(k-\ell)]$$

This is a profitable deviation iff $W^2(0,s) > W^2(s - s_1(k-\ell), s)$, or

$$(1-\delta)c\left(1-\pi^2\delta\right)(s - s_1(k-\ell)) > (1-\delta\tilde{\pi}^2)\delta V^2(s_1(k-\ell)) - (1-\tilde{\pi}^2)\delta\theta^2$$

Now, by (2.3),

$$\delta V^2(s_1(k-\ell)) = \theta^2 - (1-\delta)cs_1(k-\ell)$$

so (substituting and simplifying) this is a profitable deviation iff

$$\left(\frac{n^2 - \delta\tilde{\pi}^2}{n^2}\right) s - \delta\tilde{\pi}^2 \left(\frac{n^2-1}{n^2}\right) s_1(k-\ell) > \frac{\theta^2}{c}$$

Now, recall that $s > s_1(k)$, so writing $s = s_1(k) + \varepsilon$, for $\varepsilon > 0$, this is a profitable deviation iff

$$\left(\frac{n^2 - \delta\tilde{\pi}^2}{n^2}\right) (s_1(k) + \varepsilon) - \delta\tilde{\pi}^2 \left(\frac{n^2-1}{n^2}\right) s_1(k-\ell) > \frac{\theta^2}{c}$$

Note that if this is satisfied for $\ell = 1$, it is satisfied for all $\ell : 1 \leq \ell < k$, so the relevant inequality is:

$$\left(\frac{n^2 - \delta\tilde{\pi}^2}{n^2}\right) (s_1(k) + \varepsilon) - \delta\tilde{\pi}^2 \left(\frac{n^2-1}{n^2}\right) s_1(k-1) > \frac{\theta^2}{c}$$

and using (2.22), the inequality becomes:

$$\left(\frac{n^2 - \delta\tilde{\pi}^2}{n^2}\right) \varepsilon > 0.$$

(ii) Next, we show that $\sigma^2(s) = s - s_1(k)$ for all $s \in (s_1(k), s_1(k+1)]$. If i contributes $e = s - s_1(k)$, she gets a payoff

$$W^2(s - s_1(k), s) = \delta V^2(s_1(k)) - (1-\delta)c(s - s_1(k))$$

Suppose i contributes an amount $e \in [0, s - s_1(k))$, say $e = s - s_1(k) - \varepsilon$, with $0 < \varepsilon \leq s - s_1(k)$. This gives her a payoff

$$W^2(e,s) = \delta V^2(s_1(k) + \varepsilon) - (1-\delta)c(s - s_1(k) - \varepsilon)$$

Note that for any $s \in (s_1(k), \overline{s}_1]$,

$$V^2(s_1(k)+\varepsilon) = \tilde{\pi}^2\delta V^2(s_1(k)+\varepsilon-\sigma^2(s_1(k)+\varepsilon)) + (1-\tilde{\pi}^2)\theta^2 - \pi^2 c(1-\delta)\sigma^2(s_1(k)+\varepsilon)$$

Now, by the arguments in part (i), it is sufficient to consider $\sigma^2(s_1(k)+\varepsilon) = \varepsilon$. Therefore

$$\begin{aligned} V^2(s_1(k)+\varepsilon) &\leq \tilde{\pi}^2\delta V^2(s_1(k)) + (1-\tilde{\pi}^2)\theta^2 - \pi^2 c(1-\delta)\varepsilon \\ &= \tilde{\pi}^2\delta\theta^2 + (1-\tilde{\pi}^2)\theta^2 - \pi^2 c\left[\tilde{\pi}^2\delta(1-\delta)s_1(k) + (1-\delta)\varepsilon\right] \end{aligned}$$

and thus, for $e = s - s_1(k) - \varepsilon$ ($0 < \varepsilon \leq s - s_1(k)$):

$$W^2(e,s) \leq \delta\tilde{\pi}^2\delta\theta^2 + \delta(1-\tilde{\pi}^2)\theta^2 - \delta(1-\delta)\pi^2 c\left[\tilde{\pi}^2\delta s_1(k) + \varepsilon\right] - (1-\delta)c(s-\overline{s}_1-\varepsilon)$$

Thus contributing $s - s_1(k)$ beats contributing $s - s_1(k) - \varepsilon$ ($0 < \varepsilon \leq s - s_1(k)$) if:

$$\delta\tilde{\pi}^2\theta^2 \geq \delta\pi^2 c s_1(k)\left[1-\tilde{\pi}^2\delta\right] + c\varepsilon\left[1-\delta\pi^2\right]$$

Note the RHS is increasing in ε. Thus, this inequality holds for all $\varepsilon \in (0, s - s_1(k)]$ if it hods for $\varepsilon = s - s_1(k)$, or, rearranging, if

$$s \leq \frac{1}{c}\left(\frac{n^2}{n^2-\delta\tilde{\pi}^2}\right)\theta^2 + \delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)s_1(k) = s_1(k+1)$$

It follows that for any $s \in (s_1(k), s_1(k+1)]$, $\sigma^2(s) = s - s_1(k)$. Given this, in turn, the bound we used for $V^2(s_1(k)+\varepsilon)$ above is actually tight (so that the statement is if and only if). It follows that for all $s > s_1(k+1)$, moving the project to $s' \in (s_1(k), s_1(k+1))$ is clearly dominated by moving the project to $s_1(k+1)$, since the project is completed at the same time, with lower expected costs. Thus $\sigma^2(s) \leq s - s_1(k+1)$ for all $s > s_1(k+1)$.

An induction argument concludes the proof. □

Lemma 10. *Suppose $\overline{s}_1 < \phi_1$ and $\sigma^1(s) = s$ for all $s \leq \phi_1$. Then $\sigma^2(s) = 0$ for all $s \in [\overline{s}_1, \phi_1]$.*

Proof of Lemma 10. (i) First, note that for all $s \in [\overline{s}_1, \phi_1]$, $\sigma^2 < s - s_1(k)$ for all k in the sequence. The formal argument is the same as part (i) of the proof of Lemma 9, and therefore omitted here.

(ii) Now suppose for $s \in [\overline{s}_1, \phi_1]$, $\sigma^2(s) \in [0, s-\overline{s}_1]$, say $\sigma^2(s) = s-\overline{s}_1-\varepsilon$, with $0 \leq \varepsilon \leq s-\overline{s}_1$. This gives her a payoff

$$W^2(e,s) = \delta V^2(\overline{s}_1+\varepsilon) - (1-\delta)c(s-\overline{s}_1-\varepsilon)$$

Now, by part (i), any contribution $e \in [0, s-\overline{s}_1]$ will never result in a move to a state in which (directly or indirectly), a type 2 player advances the project to any $s_1(k)$ in the sequence. Thus $V^2(\overline{s}_1+\varepsilon) \leq (1-\tilde{\pi}^2)\theta^2$. Hence, for any $e = s-\overline{s}_1-\varepsilon$, with $0 \leq \varepsilon \leq s-\overline{s}_1$,

$$W^2(e,s) \leq \delta(1-\tilde{\pi}^2)\theta^2 - (1-\delta)c(s-\overline{s}_1-\varepsilon)$$

It follows that if $\sigma^2(s) > 0$ for any such s, a type two agent would gain by deviating to contributing zero.

To complete the proof, we show that if $\sigma^2(s) = 0$ for all $s \in [\overline{s}_1, s_1(1)]$, type 2 agents don't have a profitable deviation in this interval. Note that for any such s:

$$V^2(s) = \left(\frac{1-\tilde{\pi}^2}{1-\delta\tilde{\pi}^2}\right)\theta^2$$

A deviation to any cutpoint $s_1(k)$ gives

$$\begin{aligned} W^2(s - s_1(k), s) &= \delta V^2(s_1(k)) - c(1-\delta)(s - s_1(k)) \\ &= \theta^2 - (1-\delta)cs_1(k) - c(1-\delta)(s - s_1(k)) = \theta^2 - (1-\delta)cs \end{aligned}$$

where we used (2.3). This is a profitable deviation iff $W^2(s - s_1(k), s) > \delta V^2(s)$ or

$$s < \left(\frac{1}{1-\delta\tilde{\pi}^2}\right)\frac{\theta^2}{c} = \overline{s}_1,$$

a contradiction.

□

Lemma 11. *Suppose $w_1(1) > w_2(1)$. In the unique symmetric equilibrium,*

1. *σ^2 is a cutpoint strategy in $[0, \overline{s}_1)$ with $s_1(k)$ given by (2.6) for all $k \geq 1$ and $\lim_{k\infty} s_1(k) = \overline{s}_1 < \phi_1$.*
2. *$\sigma^2(s) = 0, \forall s \in [\overline{s}_1, \phi_1]$*
3. *$\sigma^1(s) = s, \forall s \in [0, \phi_1]$, with*

$$\phi_1 = \left(\frac{1}{1-\delta+\delta\tilde{\pi}^1\left(\frac{n^1-1}{n^1}\right)}\right)\frac{\theta^1}{c}. \tag{2.24}$$

4. *At ϕ_1, type 1 and 2 agents' equilibrium payoffs are given by:*

$$V^1(\phi_1) = \left(\frac{(1-1/n^1)\tilde{\pi}^1}{1-\delta+(1-1/n^1)\delta\tilde{\pi}^1}\right)\theta^1 \quad \textit{and} \quad V^2(\phi_1) = \left(\frac{1-\tilde{\pi}^2}{1-\delta\tilde{\pi}^2}\right)\theta^2 \tag{2.25}$$

Proof of Lemma 11. For the first part, assume towards a contradiction, that there exists a $\hat{k} < \infty$ such that

$$s_1(\hat{k}) < \hat{\phi}_1(\hat{k}) \leq s_1(\hat{k}+1).$$

where $\hat{\phi}_1(\hat{k})$ is such that $\sigma^1(s) = s, \forall s \leq \hat{\phi}_1(\hat{k})$ and $\sigma^1(s) \leq s - \hat{\phi}_1, \forall s > \hat{\phi}_1(\hat{k})$.

Step 1. Note that if $s_1(\hat{k}) < \hat{\phi}_1(\hat{k}) \leq s_1(\hat{k}+1)$, it must be that $\sigma^2(s) = s - s_1(\hat{k})$ and $\sigma^1(s) = s$ for all $s \in (s_1(k), \hat{\phi}_1(\hat{k})]$, so type 1 agent's value at $\hat{\phi}_1(\hat{k})$ is:

$$V^1(\hat{\phi}_1(\hat{k})) = \tilde{\pi}^1\theta^1 - \pi^1 c(1-\delta)\hat{\phi}_1(\hat{k}) + \tilde{\pi}^2\delta V^1(s_1(\hat{k})) \tag{2.26}$$

where $V^1(\hat{\phi}_1(\hat{k}))$ must verify:

$$\delta V^1(\hat{\phi}_1(\hat{k})) = \theta^1 - c(1-\delta)\hat{\phi}_1(\hat{k})$$

and replacing in (2.26) we get

$$\hat{\phi}_1(\hat{k}) = w^1(1)\left[1 + \frac{\delta\tilde{\pi}^2}{1-\delta}\frac{\theta^1 - \delta V^1(s_1(\hat{k}))}{\theta^1}\right] \tag{2.27}$$

Step 2. We then define[12]

$$\psi(k) := \frac{1 - \left(\delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)\right)^k}{1-\delta\tilde{\pi}^2}\frac{\theta^2}{c}.$$

Of course, for all $k \leq \hat{k}$, $s_1(k) = \psi(k)$. Thus, if $\psi(\hat{k}+1) < \hat{\phi}_1(\hat{k})$, then it *cannot* be that there exists an equilibrium which prescribes $\hat{k}$ cut-points for type-2 players below $\hat{\phi}_1(\hat{k})$. This is because while type-1 agents contribute to complete the project, the maximal amount any type-2 is willing to contribute to obtain the state $s_1(\hat{k})$ is $\psi(\hat{k}+1) - \psi(\hat{k}) < \hat{\phi}_1(\hat{k}) - s_2(\hat{k})$. Thus, at $\phi_1(\hat{k})$, only a type-1 player would be willing to contribute, contradicting the hypothesized equilibrium structure.

We now show that $\phi_1(k) > \psi(k+1)$ for all $k \geq 1$, and thus, there is no equilibrium in which both players contribute at $\phi_1(k)$.

Consider first the case $k \geq 2$. From Step 2,

$$\begin{aligned}\phi_1(k) &= \left[\frac{n^1}{n^1 - \delta\tilde{\pi}^1}\right]\frac{1}{c}\left\{\left[\frac{1-\delta\tilde{\pi}^1}{1-\delta}\right]\theta^1 - \left(\frac{\delta^2}{1-\delta}\right)\tilde{\pi}^2 V^1(s_1(k))\right\} \\ &= \frac{n^1}{n^1-\delta\tilde{\pi}^1}\cdot\frac{1}{c}\left\{\theta^1\cdot\left[\frac{1-(\delta\tilde{\pi}^2)^{k+1}}{1-\delta\tilde{\pi}^2}\right] + \theta^2 T(n^2)\right\}\end{aligned}$$

where

$$T(n^2) \equiv \frac{\delta^2\tilde{\pi}^1\tilde{\pi}^2\left(1 - (\delta\tilde{\pi}^2)^k\left(n^2 - (n^2-1)\left(1 - \left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)^k\right)\right)\right)}{n^1(1-\delta\tilde{\pi}^2)^2}$$

results from substituting $V^1(s_1(k))$ from (2.21) and simplifying. Noting that

$$\frac{\partial}{\partial n^2}\left(n^2 - (n^2-1)\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)^k\right) = 1 + k\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)^{k-1} - (k+1)\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)^k,$$

and that for any k, the function

$$f(x) = k\cdot x^{k-1} - (k+1)x^k$$

[12]Note that ψ is meant to disambiguate between the actual cutpoints $s_1(k)$, and those cutpoints of type 2 agents that would exist if there were arbitrarily many cutpoints below ϕ_1.

has its maximum at $x = \frac{k-1}{k+1}$ with $f(\frac{k-1}{k+1}) = \left(\frac{k-1}{k+1}\right)^{k-1}$, we see that

$$\frac{\partial}{\partial n^2}\left(n^2 - (n^2-1)\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)^k\right) > 0.$$

Moreover, because

$$\lim_{n^2\to\infty}\left[n^2 - (n^2-1)\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)^k\right] = (1-\delta\tilde{\pi}^2)k,$$

we see that

$$T(n^2) > \lim_{n^2\to\infty} T(n^2) = \frac{\Delta^2\tilde{\pi}^1\tilde{\pi}^2\left(1-(\delta\tilde{\pi}^2)^k\left(1+(1-\delta\tilde{\pi}^2)\cdot k\right)\right)}{n^1(1-\delta\tilde{\pi}^2)^2}.$$

Now, $(\delta\tilde{\pi}^2)^k(1+k(1-\delta\tilde{\pi}^2))$ is decreasing in k. Hence,

$$T(n^2) > \frac{\Delta^2\tilde{\pi}^1\tilde{\pi}^2\left(1-(\delta\tilde{\pi}^2)(2-\delta\tilde{\pi}^2)\right)}{n^1(1-\delta\tilde{\pi}^2)^2} > 0.$$

Therefore, for any $k \geq 2$,

$$\frac{\hat{\phi}_1(k)}{\psi(k+1)} > \frac{n^1}{n^1-\delta\tilde{\pi}^1}\cdot\frac{\theta^1}{\theta^2}\cdot\frac{1-(\delta\tilde{\pi}^2)^{k+1}}{1-\left(\delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)\right)^{k+1}}.$$

Further using the fact that $w^1(1) > w^2(1) \implies \frac{n^1}{n^1-\delta\tilde{\pi}^1}\cdot\frac{\theta^1}{\theta^2} > \frac{n^2}{n^2-\delta\tilde{\pi}^1}$, we obtain

$$\begin{aligned}
\frac{\hat{\phi}_1(k)}{\psi(k+1)} &> \frac{n^2}{n^2-\delta\tilde{\pi}^2}\cdot\frac{1-(\delta\tilde{\pi}^2)^{k+1}}{1-\left(\delta\tilde{\pi}^2\left(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2}\right)\right)^{k+1}} \\
&> n^2(n^2-\delta\tilde{\pi}^2)^k\cdot\left[\frac{1-(\delta\tilde{\pi}^2)^{k+1}}{(n^2-\delta\tilde{\pi}^2)^{k+1}-(\delta\tilde{\pi}^2)^{k+1}(n^2-1)^{k+1}}\right] \\
&> n^2(n^2-\delta\tilde{\pi}^2)^k\cdot\left[\frac{1-(\delta\tilde{\pi}^2)^{k+1}}{(1-(\delta\tilde{\pi}^2)^{k+1})(n^2-\delta\tilde{\pi}^2)^{k+1}}\right] \\
&> \frac{n^2(n^2-\delta\tilde{\pi}^2)^k}{(n^2-\delta\tilde{\pi}^2)^{k+1}} > 1,
\end{aligned}$$

the desired result.

Consider now the case where $\hat{k} = 1$. From Step 1 and Lemma 8, we have that

$$\hat{\phi}_1(1) = \frac{n^1(1+\delta\tilde{\pi}^2)}{n^1-\delta\tilde{\pi_1}}\frac{\theta^1}{c} + \frac{\delta^2\tilde{\pi}^1\tilde{\pi}^2\cdot n^2}{(n^1-\delta\tilde{\pi}^1)(n^2-\delta\tilde{\pi}^2)}\frac{\theta^2}{c}$$

by direct computation. By definition,

$$\psi(2) = \frac{1-(\delta\tilde{\pi}^2 \cdot \frac{n^2-1}{n^2-\delta\tilde{\pi}^2})^2}{1-\delta\tilde{\pi}^2}, \text{ so that}$$

$$\begin{aligned}
\frac{\hat{\phi}_1(1)}{\psi(2)} &> \frac{n^1}{n^1-\delta\tilde{\pi}^1} \cdot \frac{\theta^1}{\theta^2} \cdot \frac{1+\delta\tilde{\pi}^2}{\frac{1-(\delta\tilde{\pi}^2 \cdot \frac{n^2-1}{n^2-\delta\tilde{\pi}^2})^2}{1-\delta\tilde{\pi}^2}} \\
&> \frac{n^2}{n^2-\delta\tilde{\pi}^2} \cdot \frac{1-(\delta\tilde{\pi}^2)^2}{1-(\delta\tilde{\pi}^2 \cdot \frac{n^2-1}{n^2-\delta\tilde{\pi}^2})^2} \qquad (w_1(1) > w_2(1)) \\
&> \frac{n^2}{n^2-\delta\tilde{\pi}^2} \cdot \frac{1-(\delta\tilde{\pi}^2)^2}{(\frac{n^2-1}{n^2-\delta\tilde{\pi}^2})^2(1-(\delta\tilde{\pi}^2)^2)} \\
&= \frac{n^2(n^2-\delta\tilde{\pi}^2)^2}{(n^2-\delta\tilde{\pi}^2)(n^2-1)^2} \\
&> 1.
\end{aligned}$$

Again, we obtain the desired result in this standalone, residual case.

For the second part note that a deviation of type 2 to any $s_1(k)$ from $s \geq \overline{s}_1$ gives, using (2.3),

$$\begin{aligned}
W^2(s-s_1(k),s) &= \delta V^2(s_1(k)) - c(1-\delta)(s-s_1(k)) \\
&= \theta^2 - c(1-\delta)s
\end{aligned}$$

and this deviation is profitable if $W^2(s-s_1(k),s) > \delta V^2(s)$ or

$$\begin{aligned}
\frac{\delta\tilde{\pi}^1}{1-\delta\tilde{\pi}^2}\theta^2 &< \theta^2 - c(1-\delta)s \\
cs &< \frac{1}{1-\delta\tilde{\pi}^2}\theta^2
\end{aligned}$$

or $s < \overline{s}_1$, a contradiction.

For the third and forth part, note that given that $\overline{s}_1 < \phi_1$, and $\sigma^2(s) = 0$ for all $s \in (\overline{s}_1, \phi_1]$, we have

$$\begin{aligned}
V^1(\phi_1) &= \tilde{\pi}^1\theta^1 + \tilde{\pi}^2\delta V^1(\phi_1) - \pi^1 c(1-\delta)\phi_1 \\
V^2(\phi_1) &= \tilde{\pi}^1\theta^2 + \tilde{\pi}^2\delta V^2(\phi_1)
\end{aligned}$$

Thus we have

$$V^1(\phi_1) = \left(\frac{1}{1-\delta\tilde{\pi}^2}\right)\left\{\tilde{\pi}^1\theta^1 - \pi^1 c(1-\delta)\phi_1\right\} \tag{2.28}$$

and

$$V^2(\phi_1) = \left(\frac{1-\tilde{\pi}^2}{1-\delta\tilde{\pi}^2}\right)\theta^2 \tag{2.29}$$

Since in equilibrium type 1 must verify

$$\delta V^1(\phi_1) = \theta^1 - c(1-\delta)\phi_1$$

(2.28) directly gives (2.24). Substituting in (2.28) and (2.29) gives (2.25). Next we show that type 1 agents don't have a profitable deviation in $[0, \phi_1]$. At ϕ_1, by definition type 1 agents are indifferent between completing the project or staying put. They clearly cannot gain by moving to a non-cutpoint point, so it is enough to consider deviations to type 2 agents' cutpoints.

Recall that in equilibrium,

$$V_1^*(s_1(1)) = \theta^1 - (1-\delta)cs_1(1)$$

Suppose that instead of playing equilibrium,

Say the agent deviates and moves to some $s_1(k)$, contributing $e' = \phi_1 - s_1(k)$. By doing this, she gets a payoff:

$$W^1(e', \phi_1) = \delta V^1(s_1(k)) - (1-\delta)c(\phi_1 - s_1(k))$$

This is not a profitable deviation iff $W^1(e', \phi_1) < \theta^1 - c(1-\delta)\phi_1$ or

$$0 \geq \delta V^1(s_1(k)) + (1-\delta)cs_1(k) - \theta^1 \equiv \mathbb{L}_k^1 \tag{2.30}$$

Using now (2.23) we have that

$$\frac{\mathbb{L}_k^1}{1-\delta} = -\theta^1 + \frac{\delta}{1-\delta}\tilde{\pi}^2\mathbb{L}_{k-1}^1 - \delta\tilde{\pi}^2 cs_1(k-1) + (1-\delta\pi^1)cs_1(k)$$

and replacing $s_1(k-1)$ with the expression from (2.22) we get

$$\begin{aligned}
\frac{\mathbb{L}_k^1}{1-\delta} &= -\theta^1 + \frac{\delta}{1-\delta}\tilde{\pi}^2\mathbb{L}_{k-1}^1 + cw_2(1)\frac{n^2 - \delta\tilde{\pi}^2}{n^2-1} - cs_1(k)\left(\frac{1-\delta\tilde{\pi}^2}{n^2-1} + \delta\pi^1\right) \\
&\leq -\theta^1 + \frac{\delta}{1-\delta}\tilde{\pi}^2\mathbb{L}_{k-1}^1 + cw_2(1)\frac{n^2 - \delta\tilde{\pi}^2}{n^2-1} - cw_2(1)\left(\frac{1-\delta\tilde{\pi}^2}{n^2-1} + \delta\pi^1\right) \\
&\leq \frac{\delta}{1-\delta}\tilde{\pi}^2\mathbb{L}_{k-1}^1 + c\left(1-\delta\pi^1\right)\left(w_2(1) - w_1(1)\right)
\end{aligned}$$

and the inequality is strict if $w_2(1) < s_1(k)$. Therefore it is necessary and sufficient for type 1 not to deviate to any $s_1(k)$ from ϕ if $\mathbb{L}_1^1 \leq 0$ which is true since

$$\begin{aligned}
\mathbb{L}_1^1 &= \delta\left(\theta^1 - \pi^1 c(1-\delta)w_2(1)\right) + (1-\delta)cw_2(1) - \theta^1 \\
&= (1-\delta)c(1-\delta\pi^1)(w_2(1) - w_1(1)) < 0
\end{aligned}$$

□

Lemma 12. *For all $s > s_1(1)$ and $m \in \{1, 2\}$, $\sigma^m(s) \leq s - s_1(1)$.*

Proof of Lemma 12. Consider first type 1 agents. Note that by construction, $s_1(1)$ is the largest s at which type 1 agents would want to finish the project outright when they anticipate that type 2 agents do not contribute at s (i.e., given that $\sigma^2(s) = 0$). Since $\sigma^2(s) = 0$ minimizes free riding

incentives, it follows that $s_1(1)$ is the largest s at which type 1 agents would want to finish the project outright, for any $\sigma^2(s)$. Thus, at any $s > s_1(1)$, $\sigma^1(s) < s$. Now, we know by Lemma 6 that in any $s \in (0, s_1(1))$, a type 1 agent prefers to finish the project outright than playing $e \in [0, s)$. By the linearity of costs, it follows that for $s > s_1(1)$ a type 1 agent would weakly prefer to finish the project outright than to move the project to $s' \in (0, s_1)$. Thus it follows that at any $s > s_1(1)$, $\sigma^1(s) \leq s - s_1(1)$.

Next consider type 2 agents. Since σ^2 is a cutpoint strategy in $[0, \left(\frac{1}{1-\delta\tilde{\pi}^2}\right)\frac{\theta^2}{c})$, we know that for any k in the sequence, and any $s > s_1(k)$, $\sigma^2(s) \leq s - s_1(k)$.Thus, for all $s > s_1(1)$, $\sigma^2(s) \leq s - \left(\frac{1}{1-\delta\tilde{\pi}^2}\right)\frac{\theta^2}{c}$. Now, we know that in equilibrium $\sigma^2(s) = 0$ for all $s \in [\left(\frac{1}{1-\delta\tilde{\pi}^2}\right)\frac{\theta^2}{c}), s_1(1)]$. It follows immediately that for any $s > s_1(1)$, moving the project to $s' \in [\left(\frac{1}{1-\delta\tilde{\pi}^2}\right)\frac{\theta^2}{c}), s_1(1))$ is dominated by moving the project to $s_1(1)$, since this doesn't increase delay and reduces expected costs. Thus, for any $s > s_1(1)$, $\sigma^2(s) \leq s - s_1(1)$. □

□

2.A.3 Proof of Proposition 9

For $\tau = 1$,

$$V^1(\phi_1) = \frac{\alpha^1}{\delta}\theta^1$$

And for $2 \leq \tau \leq h_1$:

$$V^1(\phi_\tau) = \alpha^1 V^1(\phi_{\tau-1}) \Rightarrow V^1(\phi_\tau) = (\alpha^1)^\tau \frac{1}{\delta}\theta^1$$

On the other hand, for $\tau : 2 \leq \tau \leq h_1^*$:

$$\begin{aligned}\phi_\tau - \phi_{\tau-1} &= \left(\frac{1}{1 - \delta\tilde{\pi}^2 - \delta\tilde{\pi}^1/n^1}\right)\frac{\delta V^1(\phi_{\tau-1})}{c} \\ &= \phi_1(\alpha^1)^{\tau-1}\end{aligned}$$

where we used (2.24) for $\tau = 1$. Summing up over all $2 \leq \tau \leq h_1$ we get

$$\phi_{h_1} = \phi_1\left[\sum_{\tau=0}^{h_1-1}(\alpha^1)^\tau\right] = \phi_1\frac{1-(\alpha^1)^{h_1}}{(1-\alpha^1)} \tag{2.31}$$

Thus, there is no alternation of effort on the equilibrium path iff $q < \phi_{h_1}$, or

$$q < \phi_1\frac{1-(\alpha^1)^{h_1}}{1-\alpha^1} \quad (*)$$

□

2.A.4 Proof of Proposition 10

In the text we prove that if $\Delta^m < \Delta^{m'}$ then all stints of type m were one-step stints except for potentially the first one, which is the case only if $m' = 1$. We prove the result for $\Delta^2 < \Delta^1$. The case of $\Delta^2 > \Delta^1$ is analogous.

Note that for $k-1$ even, (2.15) gives

$$\Omega^2(\phi_{H(k-1)}) = \Delta^2\Omega^2(\phi_{H(k-2)}) \qquad \Delta^2 < \Omega^2(\phi_{H(k-1)}) < 1$$

so

$$\Delta^2 < \Omega^1(\phi_{H(k-2)}) < 1 \tag{2.32}$$

and for k odd, (2.15) gives

$$\Omega^1(\phi_{H(k)}) = [\Delta^1]^{h(k)}\Omega^1(\phi_{H(k-1)}) \qquad \Delta^1 < \Omega^1(\phi_{H(k)}) < 1$$

so

$$\Delta^1 < [\Delta^1]^{h(k)}\Omega^1(\phi_{H(k-1)}) < 1$$

or multiplying by Δ_2

$$\Delta^2\Delta^1 < [\Delta^1]^{h(k)}\Omega^1(\phi_{H(k-2)}) < \Delta^2 \tag{2.33}$$

Note that the left hand side of (2.33) and $\Omega^1(\phi_{H(k-2)}) < 1$ from (2.32) imply that $\Delta^2 < [\Delta^1]^{h(k)-1}$.

Now consider $k-2$ which is odd and gives

$$\Delta^1 < \Omega^1(\phi_{H(k-2)}) < 1$$

so

$$\begin{aligned}
\Delta^1 &< \Delta^2\Omega^1(\phi_{H(k-1)}) && \left(\text{by } \Omega^2(\phi_{H(k-1)}) = \Delta^2\Omega^2(\phi_{H(k-2)})\right)\\
[\Delta^1]^{h(k)+1} &< [\Delta^1]^{h(k)}\Delta^2\Omega^1(\phi_{H(k-1)}) && \left(\text{by multiplying for } [\Delta^1]^{h(k)}\right)\\
[\Delta^1]^{h(k)+1} &< \Delta^2 && \left(\text{by } [\Delta^1]^{h(k)}\Omega^1(\phi_{H(k-1)}) < 1\right)
\end{aligned}$$

Note that it must be that for all k odd

$$\begin{aligned}
[\Delta^1]^{h(k)+1} < \Delta^2 < [\Delta^1]^{h(k)-1}\\
h(k) - 1 < \frac{\log[\Delta^2]}{\log[\Delta^1]} < h(k) + 1\\
\frac{\log[\Delta^2]}{\log[\Delta^1]} - 1 < h(k) < \frac{\log[\Delta^2]}{\log[\Delta^1]} + 1
\end{aligned}$$

Assume first that $\log[\Delta^2] = \ell \log[\Delta^1]$ for some $\ell \in \mathbb{Z}^+$. Then we have that $\ell - 1 < h(k) < \ell + 1$ which implies that $h(k) = \ell$. Let now

$$\lfloor y \rfloor \equiv \sup\{x \in \mathbb{Z}_+ : x \leq y\}$$

and assume that

$$\frac{\log[\Delta^2]}{\log[\Delta^1]} = \left\lfloor \frac{\log[\Delta^2]}{\log[\Delta^1]} \right\rfloor + \epsilon$$

for some $\epsilon \in (0,1)$. It follows then that

$$\left\lfloor \frac{\log[\Delta^2]}{\log[\Delta^1]} \right\rfloor - 1 + \epsilon < h(k) < \left\lfloor \frac{\log[\Delta^2]}{\log[\Delta^1]} \right\rfloor + 1 + \epsilon$$
$$\left\lfloor \frac{\log[\Delta^2]}{\log[\Delta^1]} \right\rfloor \leq h(k) \leq \left\lfloor \frac{\log[\Delta^2]}{\log[\Delta^1]} \right\rfloor + 1$$

□

2.A.5 Proof of Theorem 3

We prove the theorem for the case of public goods with an infinite-stream of payoffs θ^m in each period and then analogize to the case of a finite stream of payoffs of length T, which allows us to make the efficiency comparison as $\delta \to 1$.

Lemma 13.

1. *Suppose* $(\Delta^1)^y < \Delta^2 < \Delta^1$ *for some* $y \in \mathbb{N}$. *Let* y^* *denote the smallest integer* y^* *such that this inequality holds. Then,*

$$\tilde{\phi} \leq \frac{\theta^1}{c(1-\delta\tilde{\pi}^2-\pi_1)}\left[\frac{1-(\alpha^1)^{h_1}}{1-\alpha^1} + \frac{(\alpha^1)^{h_1}\beta^1(1-(\alpha^1)^{y^*})}{(1-\alpha^1)(1-\beta^1(\alpha^1)^{y^*})}\right] + \frac{(\beta^2)^{h_1}\theta^2}{c(1-\delta\tilde{\pi}^1-\delta\pi_2)(1-\alpha^2(\beta^2)^{y^*})}.$$

2. *Suppose* $(\Delta^2)^y < \Delta^1 < \Delta^2$ *for some* $y \in \mathbb{N}$. *Let* y^* *denote the smallest integer* y^* *such that this inequality holds. Then,*

$$\tilde{\phi} \leq \frac{\theta^1}{c(1-\delta\tilde{\pi}^2-\pi_1)}\left[\frac{1-(\alpha^1)^{h_1}}{1-\alpha^1} + \frac{(\alpha^1)^{h_1}(\beta^1)^{y^*}}{1-\alpha^1(\beta^1)^{y^*}}\right] + \frac{(\beta^2)^{h_1}(1-(\alpha^2)^{y^*})\theta^2}{c(1-\delta\tilde{\pi}^1-\delta\pi_2)(1-\alpha^2)(1-\beta^2(\alpha^2)^{y^*})}.$$

3. *Suppose* $\Delta_1 = \Delta_2 = \Delta$. *Then*

$$\tilde{\phi} \leq \frac{\theta^1}{c(1-\delta\tilde{\pi}^2-\pi_1)}\left[\frac{1-(\alpha^1)^{h_1}}{1-\alpha^1} + \frac{(\alpha^1)^{h_1}\beta^1}{(1-\beta^1\alpha^1)}\right] + \frac{(\beta^2)^{h_1}\theta^2}{c(1-\delta\tilde{\pi}^1-\delta\pi_2)(1-\alpha^2\beta^2)}.$$

Proof of lemma 13. For the sake of brevity, we prove (1), as (2) is analogous. Part (3) follows directly from (1) or (2), making $y^* = 1$.

From (2.24), we have

$$\phi_1 = \left(\frac{1}{1-\delta\tilde{\pi}^2-\delta\pi_1}\right)\frac{\theta^1}{c}$$
$$\phi_2 = \phi_1 + \left(\frac{1}{1-\delta\tilde{\pi}^2-\delta\pi_1}\right)\frac{\delta V^1(\phi_1)}{c}$$
$$= \left(\frac{1}{1-\delta\tilde{\pi}^2-\delta\pi_1}\right)\frac{(1+\alpha^1)\theta^1}{c}$$

$$\vdots$$

$$\phi_{h_1} = \left(\frac{1}{1-\delta\tilde{\pi}^2-\delta\pi_1}\right)\frac{\left(1+\alpha^1+\ldots+\alpha^{h_1-1}\right)\theta^1}{c}$$
$$= \left(\frac{1}{1-\delta\tilde{\pi}^2-\delta\pi_1}\right)\left(\frac{1-(\alpha^1)^{h_1}}{1-\alpha^1}\right)\frac{\theta^1}{c}.$$

At ϕ_{h_1+1}, the active contributor switches from type 1 to type 2. Hence,

$$\phi_{h_1+1} = \phi_{h_1} + \left(\frac{1}{1-\delta\tilde{\pi}^1-\delta\pi_2}\right)\frac{\delta V^2(\phi_{h_1})}{c}$$
$$= \phi_{h_1} + \left(\frac{1}{1-\delta\tilde{\pi}^1-\delta\pi_2}\right)\frac{(\beta^2)^{h_1}\theta^2}{c}.$$

In the case (1) being analyzed, $H(2) = h_1 + 1$ since each stint of type 2 agents is of length exactly 1. Continuing forward,

$$\phi_{H(3)} = \phi_{H(2)} + \left(\frac{1}{1-\delta\tilde{\pi}^2-\delta\pi_1}\right)\frac{(\alpha^1)^{h_1}\beta^1\theta^1}{c}\left(\frac{1-(\alpha^1)^{h_3}}{1-\alpha^1}\right),$$

analogous to the formulation of ϕ_{h_1}. Moreover, note that by case (1), $h_3 \leq y^*$, and substituting gives us

$$\phi_{H(3)} \leq \phi_{H(1)} + \left(\frac{1}{1-\delta\tilde{\pi}^1-\delta\pi_2}\right)\frac{(\beta^2)^{h_1}\theta^2}{c} + \left(\frac{1}{1-\delta\tilde{\pi}^2-\delta\pi_1}\right)\frac{(\alpha^1)^{h_1}\beta^1\theta^1}{c}\left(\frac{1-(\alpha^1)^{y^*}}{1-\alpha^1}\right).$$

In general, note that the Lth stint for $L > 1$ odd will consist of at most y^* steps actively contributed by type 1 agents, and the $(L+1)-$th stint one step actively contributed by type 2 agents. With this emergent pattern, for any $L \geq 1$, the following formula holds for odd numbered stints:

$$\phi_{H(2L+1)} \leq \phi_{H(1)} + \left(\frac{1}{1-\delta\tilde{\pi}^1-\delta\pi_2}\right)\frac{(\beta^2)^{h_1}\theta^2}{c}\sum_{k=0}^{L-1}\left(\alpha^2(\beta^2)^{y^*}\right)^k +$$
$$\left(\frac{1}{1-\delta\tilde{\pi}^2-\delta\pi_1}\right)\frac{(\alpha^1)^{h_1}\beta^1\theta^1}{c}\left(\frac{1-(\alpha^1)^{y^*}}{1-\alpha^1}\right)\sum_{k=0}^{L-1}\left(\beta^1(\alpha^1)^{y^*}\right)^k. \quad (2.34)$$

Substituting for $\phi_{H(1)}$ from above and taking the limit as L goes to infinity in (2.34) gives us

$$\tilde{\phi} \leq \frac{\theta^1}{c(1-\delta\tilde{\pi}^2-\pi_1)}\left[\frac{1-(\alpha^1)^{h_1}}{1-\alpha^1} + \frac{(\alpha^1)^{h_1}\beta^1(1-(\alpha^1)^{y^*})}{(1-\alpha^1)(1-\beta^1(\alpha^1)^{y^*})}\right] +$$
$$\frac{(\beta^2)^{h_1}\theta^2}{c(1-\delta\tilde{\pi}^1-\delta\pi_2)(1-\alpha^2(\beta^2)^{y^*})}.$$

□

Suppose there exists minimal positive integer y^* such that $(\Delta^1)^{y^*} < \Delta^2 < \Delta^1$. From Lemma 13 (1),

$$\begin{aligned}\tilde{\phi} &\leq \frac{\theta^1}{c(1-\delta\tilde{\pi}^2-\delta\pi_1)}\left[\frac{1}{1-\alpha^1}-\frac{(\alpha^1)^{h_1}(1-\beta^1)}{(1-\alpha^1)(1-\beta^1(\alpha^1)^{y^*})}\right]+\frac{(\beta^2)^{h_1}\theta^2}{c(1-\delta\tilde{\pi}^1-\delta\pi_2)(1-\alpha^2(\beta^2)^{y^*})}\\ &< \frac{\theta^1}{c(1-\delta\tilde{\pi}^2-\delta\pi_1)(1-\alpha^1)}+\frac{\theta^2}{c(1-\delta\tilde{\pi}^1-\delta\pi_2)(1-\alpha^2)},\end{aligned}$$

where the second inequality follows from the fact that $\beta^2 < 1$. Furthermore, for each type $j \in \{1, 2\}$,

$$1-\alpha_j = 1-\frac{(1-1/n_j)\tilde{\pi}_j\delta}{1-\delta\tilde{\pi}_{-j}-\delta\pi_j} = \frac{1-\delta}{1-\delta\tilde{\pi}_j-\delta\pi_j}$$

Substituting into the above inequality, and re-arranging we obtain

$$c\tilde{\phi} < \frac{\theta^1+\theta^2}{1-\delta}$$

for any fixed $\delta < 1$, in the case of infinte-lasting flow payoffs. For a project whose flow payoffs last exactly T periods, the analogous expression is

$$c\tilde{\phi} < \frac{(1-\delta^T)(\theta^1+\theta^2)}{1-\delta} = \hat{\theta}^1+\hat{\theta}^2.$$

Otherwise, there exists a positive integer y^* such that $(\Delta^2)^{y^*} < \Delta^1 < \Delta^2$. From Lemma 13 (2),

$$\begin{aligned}\tilde{\phi} &\leq \frac{\theta^1}{c(1-\delta\tilde{\pi}^2-\delta\pi_1)}\left[\frac{1-(\alpha^1)^{h_1}}{1-\alpha^1}+\frac{(\alpha^1)^{h_1}(\beta^1)^{y^*}}{1-\alpha^1(\beta^1)^{y^*}}\right]+\frac{(\beta^2)^{h_1}(1-(\alpha^2)^{y^*})\theta^2}{c(1-\delta\tilde{\pi}^1-\delta\pi_2)(1-\alpha^2)(1-\beta^2(\alpha^2)^{y^*})}\\ &< \frac{\theta^1}{c(1-\delta\tilde{\pi}^2-\delta\pi_1)(1-\alpha^1)}+\frac{\theta^2}{c(1-\delta\tilde{\pi}^1-\delta\pi_2)(1-\alpha^2)},\end{aligned}$$

so that we again obtain the first desired result

$$c\tilde{\phi} \leq \frac{\theta^1+\theta^2}{1-\delta}.$$

Redacting the flow payoffs from the public good to a horizon of length T gives the analogous expression

$$c\tilde{\phi} \leq \hat{\theta}^1+\hat{\theta}^2.$$

Since for any $\delta < 1$, we have $c\tilde{\phi} < \hat{\theta}^1+\hat{\theta}^2$, then for $n^m > 1$ for some $m \in \{1, 2\}$,

$$\lim_{\delta\to 1} c\tilde{\phi} \leq \hat{\theta}^1+\hat{\theta}^2 < n^1\hat{\theta}^1+n^2\hat{\theta}^2.$$

Since $\tilde{\phi}$ is the maximal quantity q corresponding to projects completed in equilibrium, this gives us the second result.

□

2.A.6 Proof of Theorem 11

We show that the expected completion time of a project that requires L contribution stints, $\{h_1, \ldots, h_L\}$, is given by

$$\mathcal{E}\left(\{h_\ell\}_{\ell=1}^{L}\right) = \sum_{\ell=1}^{L} (\tilde{\pi}_{m_\ell})^{h_\ell - 1} \left[h_\ell + \frac{\tilde{\pi}_{m_{-\ell}}}{\tilde{\pi}_{m_\ell}}\right],$$

where $m_\ell = 1$ if ℓ is odd and 2 if ℓ is even, and $m_{-\ell} = m \in M \neq m_\ell$. The second part of the Theorem follows by the argument in the text.

Consider the sequence $\{h_\ell\}_{\ell=1}^{L}$ and define the random variable $x\left(\{h_\ell\}_{\ell=1}^{L}\right)$, as the number of periods it takes to finish the project. Note that this random variable has the following distribution:

$$\begin{aligned}
x\left(\{h_\ell\}_{\ell=1}^{L}\right) &= h_L + x\left(\{h_\ell\}_{\ell=1}^{L-1}\right) && \text{w.p.} \quad (\tilde{\pi}^1)^{h_L} \\
x\left(\{h_\ell\}_{\ell=1}^{L}\right) &= h_L + 1 + x\left(\{h_\ell\}_{\ell=1}^{L-1}\right) && \text{w.p.} \quad (\tilde{\pi}^1)^{h_L}\tilde{\pi}^2 \\
&\cdots \\
x\left(\{h_\ell\}_{\ell=1}^{L}\right) &= h_L + k + x\left(\{h_\ell\}_{\ell=1}^{L-1}\right) && \text{w.p.} \quad (\tilde{\pi}^1)^{h_L}(\tilde{\pi}^2)^k
\end{aligned}$$

It follows that

$$x\left(\{h_\ell\}_{\ell=1}^{L}\right) = x\left(h_L\right) + x\left(\{h_\ell\}_{\ell=1}^{L-1}\right) = \sum_{\ell=1}^{L} x\left(h_\ell\right)$$

and therefore

$$\mathbb{E}\left(x\left(\{h_\ell\}_{\ell=1}^{L}\right)\right) = \sum_{\ell=1}^{L} \mathbb{E}\left(x\left(h_\ell\right)\right)$$

Note that, letting $m_\ell = 1$ if ℓ is odd and 2 if ℓ is even, and $m_{-\ell} = m \in M \neq m_\ell$, we have:

$$\begin{aligned}
\mathbb{E}\left(x\left(h_\ell\right)\right) &= (\tilde{\pi}_{m_\ell})^{h_\ell} h_\ell + (\tilde{\pi}_{m_\ell})^{h_\ell}\tilde{\pi}_{m_{-\ell}}(h_\ell + 1) + \ldots \\
&\qquad + (\tilde{\pi}_{m_\ell})^{h_\ell}(\tilde{\pi}_{m_{-\ell}})^k(h_\ell + k) + \ldots \\
&= (\tilde{\pi}_{m_\ell})^{h_\ell} h_\ell + \sum_{k=1}^{\infty} (\tilde{\pi}_{m_\ell})^{h_\ell}(\tilde{\pi}_{m_{-\ell}})^k(h_\ell + k) \\
&= (\tilde{\pi}_{m_\ell})^{h_\ell}\left[h_\ell + \sum_{k=1}^{\infty} (\tilde{\pi}_{m_{-\ell}})^k(h_\ell + k)\right] \\
&= (\tilde{\pi}_{m_\ell})^{h_\ell}\left[h_\ell \sum_{k=0}^{\infty} (\tilde{\pi}_{m_{-\ell}})^k + \tilde{\pi}_{m_{-\ell}} \sum_{k=1}^{\infty} (\tilde{\pi}_{m_{-\ell}})^{k-1} k\right] \\
&= (\tilde{\pi}_{m_\ell})^{h_\ell}\left[\frac{h_\ell}{1 - \tilde{\pi}_{m_{-\ell}}} + \tilde{\pi}_{m_{-\ell}} \frac{d\sum_{k=0}^{\infty} (\tilde{\pi}_{m_{-\ell}})^k}{d\tilde{\pi}_{m_{-\ell}}}\right]
\end{aligned}$$

$$
\begin{aligned}
&= (\tilde{\pi}_{m_\ell})^{h_\ell} \left[\frac{h_\ell}{1-\tilde{\pi}_{m_{-\ell}}} + \tilde{\pi}_{m_{-\ell}} \frac{d\left(\frac{1}{1-\tilde{\pi}_{m_{-\ell}}}\right)}{d\tilde{\pi}_{m_{-\ell}}} \right] \\
&= (\tilde{\pi}_{m_\ell})^{h_\ell} \left[\frac{h_\ell}{1-\tilde{\pi}_{m_{-\ell}}} + \frac{\tilde{\pi}_{m_{-\ell}}}{(1-\tilde{\pi}_{m_{-\ell}})^2} \right] \\
&= (\tilde{\pi}_{m_\ell})^{h_\ell - 1} \left[h_\ell + \frac{\tilde{\pi}_{m_{-\ell}}}{\tilde{\pi}_{m_\ell}} \right],
\end{aligned}
$$

and therefore

$$
\mathbb{E}\left(x\left(\{h_\ell\}_{\ell=1}^{L}\right)\right) = \sum_{\ell=1}^{L} (\tilde{\pi}_{m_\ell})^{h_\ell - 1} \left[h_\ell + \frac{\tilde{\pi}_{m_{-\ell}}}{\tilde{\pi}_{m_\ell}} \right]
$$

□

Chapter 3

When Incumbency Limits Productivity*

3.1 Introduction

I consider a class of problems where in each of several periods, two agents cooperate to advance a public good, or joint project. At any given point of time, one agent, the *project leader* proposes an advancement which must be approved by her counterpart, the *respondent*, in order to be implemented by the project leader at her own cost. Projects at any interim stage of the game can be described by two fundamental characteristics, over which the agents have different preferences. Hence, agents who contribute to the project at personal cost have the opportunity to not only improve upon the project, but to also determine the *direction* in which the project moves, subject to their counterpart's approval, thus negating free-riding incentives.

There are many situations arising in economic, political, and other contexts that are encompassed by this game description. For instance, a team of scholars collaborating on a joint project makes gradual progress on its research only when all team members agree to incorporate advances made by any single member. Members of a committee (e.g., intergovernmental alliance, legislature, board of directors) often designate a leader to develop and propose improvements on existing solutions to common challenges, which must be agreed upon by all prior to implementation. Workers at a firm collaborate on a collective task, with the expertise of—and thereby assent of— each needed for

*I thank my adviser Leeat Yariv for her patient guidance of this work, and Alessandro Lizzeri for his encouragement and suggestions for applications of this model in public economics. I am also grateful to Matias Iaryczower, Christopher Sandmann, Mu Zhang, Evgenii Safonov, and Pellümb Reshidi for their helpful comments. I also thank the audience at the Princeton micro-economic theory workshop for their suggestions.

the project leader to execute her vision. In each of these scenarios, agents typically differ in which solutions they prefer to implement for the project; hence, unlike in many collective action problems, there is *an inherent value to being the project leader*.

In this paper, I explore how changing the *rate of incumbency*— the probability that any given project leader retains charge of the project in the next period— affects the *productivity* of agents. Here, I define a project leader's productivity by the proportion of remaining work on the project that is completed by her. An increased probability of turnover (decreased incumbency) can affect productivity in two opposite ways: first, it can incentivize a project leader to be *more productive*, as she anticipates that she may lose the ability to shape the project in the next period. By exerting more effort in the present, she can permanently alter the project to suit her interests before turnover occurs. On the other had, decreased incumbency also means the respondent anticipates that she will become the ensuing project leader with increased probability. Thus, she has an incentive to be more demanding in the present, which reduces the per-unit benefit to the current project leader and may lead her to be *less productive*. The trade-offs are, of course, exactly reversed when incumbency increases.

We examine these trade-offs in a model in which the two agents differ in their costs of contributing to the project. One agent is *efficient* in the sense that she can advance an equivalent amount of progress at a lower cost than her counterpart. The central result (Proposition 14) of the paper is that *increased incumbency decreases the productivity of an efficient project leader and increases that of a relatively inefficient one.*

Why does this occur? First, an efficient project leader is always able to generate more surplus in the game since her costs are lower. Because the incumbency rate is common across proposers— regardless of type— each project leader expects to retain proposership for the same proportion of future periods within the game. As a result, conditional on a particular, common role (project leader or responder), the payoff of the efficient agent is always higher than that of the inefficient one. Now consider the trade-offs discussed in the preceding paragraphs. For the efficient project leader, the gain in payoff from increased incumbency is *greater* than the loss to her counterpart. Hence, the

increase in her leverage over the respondent (which incentivizes productivity) are dwarfed by the increase in her own future payoff, which disincentivizes productivity in the current period. On the other hand, for the inefficient project leader, the gain in leverage over the respondent— the efficient contributor— is larger than her own gains from future proposership. Hence, for her, productivity is incentivized.

3.1.1 Literature Review

The central question of this paper is examined in a variety of different settings and literatures. For instance, in public economics, there is a large empirical and theoretical literature on the effect of government turnover on public spending and debt— e.g., (Persson and Svensson, 1989), (Alesina and Tabellini, 1990), (Battaglini and Coate, 2008), (Yared, 2019). While my model is not geared specifically to this legislative environment, we can interpret the project leader as an incumbent and the respondent as the opposition, with the joint project being instituting a public good; progress requires public spending, which partially forecloses further public spending in the future. Agents' preferences diverge over the characteristics of the public good. We discuss the relationship of our main result to this literature further in Section 3.4.2.

This main question is also related to a line of inquiry in the work and organizational studies literature that analyzes how job (in)security affects worker productivity. Many of the papers in the literature— e.g., (Holmström, 1999), (Kuvalekar and Lipnowski, 2020) — model this as an agency problem where misalignment in the incentives of managers and workers further distorts the incentives for worker productivity. The principal-agent setting is often appropriate; the present model thus comments on how team productivity might be altered within *cooperative work settings*. Here, we can interpret job insecurity as not "being fired," but rather ceding control of a firm-wide project to another fellow employee.

The analytical framework of this paper is closely related to (Acharya and Ortner, 2021), which similarly considers a game of dynamic cooperation in which agents differ in their preferences over which of two characteristics a joint project exhibits. In particular, my model draws its recursivity

from their approach. (Acharya and Ortner, 2021) studies dynamic search within bargaining— in each period, a random alternative is drawn from the set of Pareto improvements on the status quo. Two agents with divergent interests must both agree in order to implement it. The focus of the paper, in contrast to our own, is on bargaining frictions and the path of implemented outcomes.

More broadly, the paper sits within both the literature on dynamic bargaining with an endogenous status quo as well as the literature on contributions to public goods. Like in the former, our agents enact irreversible “policies" that alter the incentives for enacting future outcomes. In particular, in both our paper and this literature, the dynamic linkage between agreements distorts the incentives of agents. Papers that examine the consequences of this broad class of incentives include (Baron, 1996), (Duggan and Kalandrakis, 2012), (Bowen et al., 2014), and (Dziuda and Loeper, 2016), among others, with (Eraslan et al., 2020) providing an excellent recent survey.

Finally, this paper also is adjacent to much of the public goods literature following (Admati and Perry, 1991)— for example, (Compte and Jehiel, 2004), (Kessing, 2007), and (Iaryczower et al., 2021). Similar to my model, these papers consider an environment in which contributions are made to a joint project gradually by several agents over many periods. In general, agents rotate in the role of contributing, which generates a free rider problem. Recent work in this literature also consider the possibility of differentiating contributions to suit one’s own preferences, like in the present model, which enables rent extraction. Among these, (Bonatti and Rantakari, 2016) is perhaps most closely related to our model; here, the authors consider two agents competing to complete a “proposal" which can be differentiated along an ideological parameter. In a second stage of the game, a committee of agents also votes on whether to implement a completed proposal or wait until the second proposal is presented. Hence, like our model, (Bonatti and Rantakari, 2016) combines aspects of contribution and bargaining games.

3.1.2 Paper Outline

The remainder of the paper is structured as follows— Section 3.2 presents the formal model of the two-agent contribution game. Next, Section 3.3 presents the basic analytic framework needed

for equilibrium characterization. We focus in Section 3.3 on equilibria in the game with finite deadlines— wherein agents can only contribute to the project for a limited number of periods. I show that there is a unique subgame perfect equilibrium (SPE) of the game with any arbitrary finite deadline which is Markov in the existing progress on the project and which exhibits a particular form of stationarity (Proposition 12).

This fact is a useful segue for presenting our main results in Section 3.4. I show first that the sequence of unique SPE for each finite deadline T converges to the unique stationary Markov perfect equilibrium of the infinite deadline game (Proposition 13). This provides a motivation for focusing on this form of equilibrium for our results. I then show the main result of the paper holds within this equilibrium— *incumbency offers the opposite productivity incentives for efficient and inefficient project leaders*. Finally, Section 3.5 offers concluding remarks.

3.2 Model

Two agents, 1 and 2, take turns to contribute to a joint project, over many periods during which advancements on the project can only be implemented with bipartisan approval. At any point of time, the status of the project is entirely captured by two distinct *characteristics*, $\mathcal{C}_1$ and $\mathcal{C}_2$, which are rival and alternative features the project can exhibit. For instance, if politicians are developing and then implementing solutions to further criminal justice reform, they can focus on alternatives that promote either deterrence or rehabilitation. Alternatives for the project are also characterized by how *efficient* (*complete*) they are. We assume that there is a fully efficient frontier, which allows us to normalize the space of alternatives. In particular, all feasible alternatives for the project are characterized by the set

$$X := \{x \in \mathbb{R}^2_+ : x_1 + x_2 \leq 1\}.$$

We let $\theta_x := x_1 + x_2 \leq 1$ describe the efficiency of alternative x. For each agent i, the utility from a given alternative x is

$$v_i(x) = x_i.^1$$

Moreover, we will impose that once an alternative $x \in X$ is implemented, the extent of either characteristic it exhibits becomes a permanent feature of the project. Implementing a new alternative thus entails an irreversible advance in solving the agents' collective action problem. An alternative, once implemented, is denoted the *status quo* $s \in X$.

3.2.1 Timeline

At the start of each period of play, agents inherit a particular $s \in X$ (implemented in the previous period). One of the two agents, $p \in \{1, 2\}$, is chosen as the *project leader* and is called upon to *advance* a new alternative to the status quo. Let

$$X(s) := \{x \mid x_i \geq s_i \; \forall \; i\}$$

represent the set of feasible Pareto improvements on the status quo, s. Agent p's *proposal* for the project at s is thus a choice $x_p \in X(s)$. Once presented with x_p, the non-proposing agent, the *respondent*, r, can either accept $(a = 1)$ or reject $(a = 0)$ it.

If a proposal is rejected, it is abandoned and not pursued. Otherwise, the project leader must source and then implement the proposed alternative. She does so by undertaking some costly action: for instance, engaging in or funding research that discovers new technologies, exerting time and

[1]Note that this model is compatible with scenarios in which there is no possibilities frontier, but the utilities of agents is asymptotic. Define $X^o := \{x \in \mathbb{R}+^2 : x_1 + x_2 < 1\}$. Consider for example the map $f : X^o \to \mathbb{R}^2_+$ where

$$f(x) = \big(f_1(x_1), f_2(x_2)\big) = \left(\frac{x_1}{1 - x_1} \quad , \quad \frac{x_2}{1 - x_2}\right).$$

Note that X^0 is isomorphic to $\mathbb{R}^2_+$ under this map and that we can therefore define $\mathbb{R}^2_+$ as the space of alternatives with $v_i(y) = f_i^{-1}(y_i)$ for any $y \in \mathbb{R}^2_+$. It is moreover sufficient for f to only be defined for X^0, for given the form of the costs c_i, the points on the boundary of X are never arrived at in any equilibrium.

effort to execute on a common project, expending “political capital" and other resources in drafting legislation, or by simply investing irreversibly in technologies or goods.

The cost to p of sourcing alternative $x \in X(s)$ is denoted $\hat{c}_p(x, s)$ and satisfies the identity

$$\frac{\hat{c}_p(x,s)}{1-\theta_s} = c_p\left(\frac{\theta_{x-s}}{1-\theta_s}\right), \tag{3.1}$$

where $c_p : [0,1] \to \overline{\mathbb{R}}_+$ is given by an increasing, differentiable, convex function such that

$$c_p(0) = 0, \quad c_p'(0) \leq 1, \quad c_p(1) = \infty.$$

That is, the cost of advancing a given alternative x depends on the current status quo s; in particular, the cost depends on and is increasing in the amount of *remaining* inefficiency $1 - \theta_s$ at s that would be erased by implementing x. This conceptualization captures the natural idea that as society advances closer to the frontier, the same *absolute gains* in efficiency are more costly. We will assume that agents have heterogeneous costs of contribution—

$$c_1 = \kappa \cdot c_2, \qquad \text{where } \kappa < 1$$

— so that agent 1 is the *efficient project leader*.

Finally, utilities are realized at the end of each period of play. In period t, agents inherit status quo s_{t-1} and implement new status quo s_t. Hence, the single-period utility for the period-t project leader p is given by

$$u_p(s_t, s_{t-1}) = (s_t)_p - \hat{c}_p(s_t, s_{t-1});$$

whereas the single-period utility for the period-t responder r is

$$u_r(s_t, s_{t-1}) = (s_t)_r.$$

We assume that agents have a (possibly infinite) *deadline* $T \in \overline{\mathbb{N}} := \{1, 2, \ldots, \infty\}$ to contribute to the project. After time T, benefits from existing projects continue to accrue to agents across an infinite horizon. That is, $s_t = s_{t+1}$ for each $t \geq T$ and the total utility of agent i is given by

$$U^i(s, k_i) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_{k_i(t)}(s_t, s_{t-1})$$

for common discount rate $\delta \in (0, 1)$. Here, $s := \{s_t\}_{t=0}^{\infty}$, where $s_0 = 0$; and, $k_i(t) \in \{p, r\}$ is the role agent i takes on in period t, with $k_i := \{k_i(t)\}_{t=1}^{\infty}$.[2] The form of our utility function mimics several situations that emerge from team settings— for instance, two scholars collaborating on a paper they wish to submit to a conference have a finite amount of time to complete it, while its reputational and innovational benefits last throughout their academic careers. We will particularly focus on the $T = \infty$ case, which better captures progress on endemic, societal projects, where deadlines are less pressing.

Finally, we assume that the *incumbency rule* π in the model is *Markov* in that it only depends on the identity of the previous proposer, and is symmetric across proposers. We let $\pi \in [0, 1]$ represent the probability that any proposer will remain in leadership from one period to the next.

3.2.2 Strategies and Equilibrium

The history at the start of time t is $h_t = \{(p_\tau, x_\tau, a_\tau)\}_{\tau<t}$, where $p_\tau \in \{1, 2\}$ is the time-τ project leader, x_τ the alternative she proposes, and a_τ the decision made by the time-τ respondent. Let h_1 therefore be any null history. Given a history h_t, we can recursively determine the status quo

$$s_\tau(h_t) = a_\tau x_\tau + (1 - a_\tau) s_{\tau-1}(h_t)$$

for each $1 \leq \tau < t$, with $s_0 = 0$. Let the set of time-t histories be denoted H_t. A pure strategy σ_i is a tuple (x_i, a_i), where

$$x_i : H_t \times \overline{\mathbb{N}} \to X$$

[2]Note that we can let $k_i(t)$ be arbitrary for any $t > T$ since the utilities of a given agent in either role is the same.

describes the proposal agent i makes as the proposer, given any history h_t and deadline T, and

$$a_i : H_t \times X \times \overline{\mathbb{N}} \to \{0, 1\}$$

her decision to accept or reject any given proposal from a time-t history with deadline T, when the respondent. For any *subgame perfect equilibrium* (SPE) $\sigma = (\sigma_1, \sigma_2)$, and each history h_t, we let $V_i^\sigma(h_t; T)$ be player i's continuation value from the (sub)game with deadline T (T more periods of play remaining) under profile σ, at h_t.

Because the respondent's action profile is binary over each history–alternative pair (h_t, x_t), we can represent an action of the respondent as an *acceptability set*

$$A_r^\sigma(h_t, T) = \{x \in X(s_{t-1}(h_t)) \;\mid\; a_r(h_t, x, T) = 1\}.$$

Let the set of all acceptability sets at h_t be $\mathcal{A}(h_t) := 2^{X(s_{t-1}(h^t))}$. We define also two refinements of subgame perfect equilibrium which will be useful in describing equilibria of the game—

Definition 13. *We call a subgame perfect equilibrium of the game σ with deadline T a Markov perfect equilibrium (MPE) if for each $i \in \{1, 2\}$,*

$$x_i(h_t, T) = x_i(h_t', T) \qquad \textit{and} \qquad a_i(h_t, x, T) = a_i(h_t', x, T)$$

for any histories h_t, h_t' for which $s_{t-1}(h_t) = s_{t-1}(h_t')$ and any $x \in X(s_{t-1}(h_t))$. Define for any status quo s, and $x \in X(s)$,

$$P_s(z) := \left(\frac{z_1 - s_1}{1 - \theta_{z-s}} \;\;,\;\; \frac{z_2 - s_2}{1 - \theta_{z-s}} \right)$$

Then such an equilibrium is, moreover, stationary Markov perfect if at any two histories h_t, $h_t' \in H_t$,

$$P_{s_{t-1}(h_t')}(x_i(h_t', T)) = P_{s_{t-1}(h_t)}(x_i(h_t, T)), \textit{ and}$$

$$P_{s_{t-1}(h'_t)}(A_i^\sigma(h'_t, T)) = P_{s_{t-1}(h_t)}(A_i^\sigma(h_t, T)).$$

Hence, strategies in Markov perfect equilibria depend only on the state variable s, the previous status quo policy implemented, and not on the history of play itself. Such equilibria, are moreover, stationary, if they mandate the same strategy— up to scaling with the size of the progress made on the project through the status quo s— from every subgame.

3.3 Preliminaries and the Finite Deadline Game

In this section, I outline results that help to motivate a *recursive structure* to equilibrium, in general. This provides a helpful prelude to the analysis required for the main result in Section 3.4. I do so by first analyzing the game with a finite deadline. In particular, I demonstrate that each finite deadline game has a unique SPE σ which is a stationary MPE. I also show that the limit as $T \to \infty$ of this sequence of equilibria exists and is the unique stationary MPE of the infinite game. This provides a foundation for focusing on stationary MPE of the infinite deadline game, as the behavior is locally generalizable to SPE in games with large T.

3.3.1 Analytic Framework

Observe that in any SPE σ, $V_i^\sigma(h_t, 0) = (s_{t-1}(h_t))_i$ for any h_t and that the $T-$ cutoff continuation value to i from implementing any alternative $z \in X(s_{t-1}(h_t))$ (net of any costs) at history h_t is

$$(1-\delta)z_i + \delta V_i^\sigma(h', T-1), \tag{3.2}$$

where $h' = h_t \cup \{(p_t, z, 1)\}$. Let also $h'' = h_t \cup \{(p_t, z, 0)\}$ The strategies σ_i therefore satisfy

$$x_i(h_t, T) \in \text{argmax}_{x \in A_{-i}^\sigma(h_t,T)} \{(1-\delta)(x_i - \hat{c}_i(x, s_{t-1}(h_t))) + \delta V_i^\sigma(h', T-1)\}$$

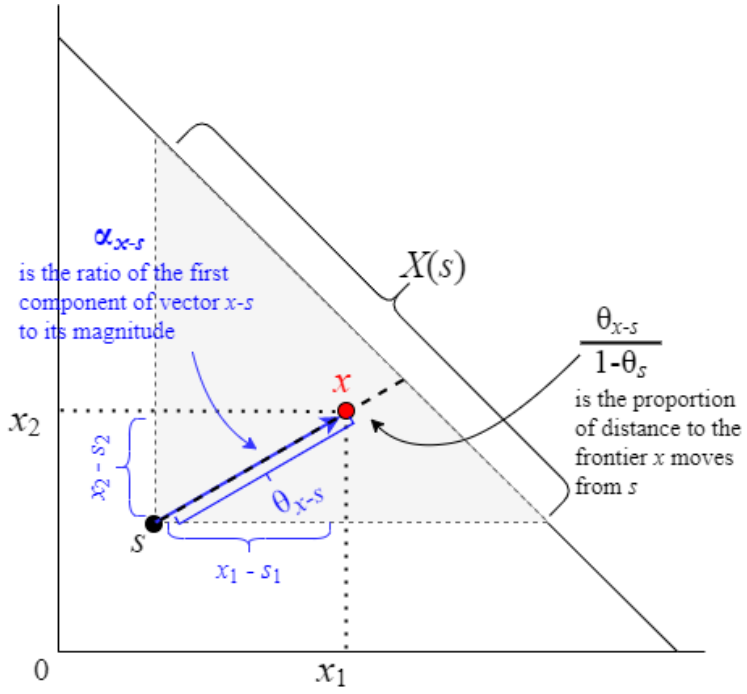


Figure 3.3.1: The policy space and bijection $(x_1, x_2) \mapsto_s \left(\alpha_{x-s}, \frac{\theta_{x-s}}{1-\theta_s}\right)$.

$$x \in A_i^\sigma(h_t, T) \Leftrightarrow (1-\delta)x_i + \delta V_i^\sigma(h', T-1) \geq (1-\delta)(s_{t-1}(h_t))_i + \delta V_i^\sigma(h'', T-1). \tag{3.3}$$

when there are T periods remaining in gameplay, at any history h_t of the game.

Define further $V_i^{j,\sigma}(h_t, T)$ be the continuation value to i at history h_t with current leader j and $V^{j,\sigma}(h_t, T) := \sum_{i=1}^2 V_i^{j,\sigma}(h_t, T)$. We let $W_i^\sigma(T) := V_i^\sigma(h_1, T)$, $W_i^{j,\sigma}(T) := V_i^j(h_1, T)$, and $W^{j,\sigma}(T) := \sum_{i=1}^2 W_i^{j,\sigma}(T)$ be the "zero-history" analogues of these value functions.

We next introduce a convenient transformation that simplifies our analysis. For any x, let $\alpha_x := \frac{x_1}{\theta_x} \in [0,1]$ denote how much alternative x favors characteristic $\mathcal{C}_1$.[3] Further define $\beta_{x,s} := \frac{\theta_{x-s}}{1-\theta_s} \in [0,1]$ as the proportion of progress remaining on the project that implementing alternative x would accomplish. Given a status quo s, consider the following bijection of any alternative $x \in X(s)$,

$$(x_1, x_2) \mapsto_s \left(\alpha_{x-s}, \frac{\theta_{x-s}}{1-\theta_s}\right).$$

[3]Since α_x is not well-defined for $x = 0$, let $\alpha_0 = 0$.

Note that $(\alpha, \beta)-$ space is isomorphic to X under the mapping

$$(\alpha, \beta) \mapsto (\alpha\beta, (1-\alpha)\beta)\,,$$

and that

$$(x_1, x_2) = (s_1 + (1-\theta_s)\alpha\beta, s_2 + (1-\theta_s)(1-\alpha)\beta)\,.$$

With some abuse of notation, we will therefore describe the strategies of players in (α, β)- terms, rather than using the standard Cartesian coordinates, within X.

Observation 1. *To see that this is, indeed, convenient, note that the cost to p from advancing a given alternative x from status quo s is simply $c_p(\beta_{x,s})$. Moreover, under Definition 13, an equilibrium is stationary Markov perfect if for each history h_t and $i \in \{1, 2\}$,*

$$\beta_{x_i(h_t,T),s_{t-1}(h_t)} = \beta_i(T) \quad \textit{and}$$
$$\left\{(\alpha, \beta) : (\alpha, \beta) = (\alpha_{x-s_{t-1}(h_t)}, \beta_{x,s_{t-1}(h_t)}) \textit{ for some } x \in A_i^\sigma(h_t, T)\right\} = A_i(T)$$

for some fixed $\beta_i(T)$ and $A_i(T)$. Thus, in any stationary MPE, the strategies of players are history-invariant in (α, β)- space.

3.3.2 Equilibrium Characterization

We proceed by analyzing first the case $T = 1$, which provides the tenor of the analysis for any finite deadline. For shorthand, let s be the current status quo $(s_{t-1}(h))$ corresponding to history h. Because this is the final period before the deadline is reached, the utility to i of implementing alternative $z \in X(s)$ is

$$(1-\delta)z_i + \delta V_i^\sigma(h', 0) = z_i.$$

Therefore, the condition for $z \in A_i^\sigma(h)$ in any SPE σ is simply $z_i \geq s_i$. Hence, $A_i^\sigma(h, 1) = X(s)$.

To fix ideas, let the project leader be 1 without loss of generality. We write $(\alpha_1(h,1), \beta_1(h,1)) = (\alpha_{x_1(h,1)-s}, \beta_{x_1(h,1),s})$ for the proposal advanced by 1 under σ. Since $A_2^\sigma(h,1) = X(s)$, and the cost $(1-\theta_s)c(\beta_1)$ of advancing (α_1, β_1) depends only on β_1, it follows that $\alpha_1(h,1) = 1$. That is, for any β_1, the surplus extracted by 1 will be maximal since it is within the acceptance set of 2. The maximization problem of the project leader is therefore given by

$$\begin{aligned} &\max_{\beta_1 \in [0,1]} [s + (1-\theta_s)(\beta_1 - (1-\delta)c_1(\beta_1))] \\ \simeq &\max_{\beta_1 \in [0,1]} [\beta_1 - (1-\delta)c_1(\beta_1)], \end{aligned}$$

which has solution

$$\beta_1(h,1) = [c_1']^{-1}\left(\frac{1}{1-\delta}\right). \tag{3.4}$$

Importantly, the strategy of the project leader is *independent of the status quo s and history h*. Note that this is also true of the responder since she accepts any proposal in (α, β)-space regardless of the history h. The analysis of the case where 2 is the project leader is, obviously, exactly analogous. From Observation 1, we thus see that (i) *there is a unique SPE of the $T = 1$ deadline game which is stationary Markov perfect* and utilizes the particular strategies solved for above.

In fact, we will drop h as an argument of this maximizer, for convenience. This stationary property allows us to rewrite the continuation values at h of both agents when 1 is the leader, in the unique SPE σ of the one-period game,

$$\begin{aligned} (V_1^{\sigma,1}(h,1), V_2^{\sigma,1}(h,1)) &= (s_1 + (1-\theta_s)\left[\beta_1(1) - (1-\delta)c_1(\beta_1(1))\right]\ ,\ s_2), \text{ as} \\ &= (s_1 + (1-\theta_s)W_1^{\sigma,1}(1)\ ,\ s_2 + (1-\theta_s)W_2^{\sigma,1}(1)). \end{aligned} \tag{3.5}$$

The second inequality follows from substituting $s = 0$ into the first equation and noting that $W_1^{\sigma,1}(1) = \beta_1(1) - (1-\delta)c_1(\beta_1(1))$, $W_2^{\sigma,1}(1) = 0$. The analogous statement also holds for $p = 2$.

We take this opportunity to simplify notation— for any *Markov perfect equilibrium* σ, we write

$$\hat{V}_i^{\sigma,j}(s,T) := V_i^{\sigma}(h_t,T) \text{ for any } h_t \text{ such that } p_{t-1}(h_t) = j,\ s_{t-1}(h_t) = s,$$
$$\hat{V}^{\sigma,j} := \sum_{i=1}^{2} \hat{V}_i^{\sigma,j}(s,T);$$

and, $\hat{W}_i^{\sigma,j}(T) := \hat{V}_i^{\sigma,j}(0,T)$, $\hat{W}^{\sigma,j}(T) := \sum_{i=1}^{j} \hat{V}_i^{\sigma,j}(0,T)$. This is possible because $V_i^{\sigma}(h_t,T)$ only depends on h_t through the identity of the *previous incumbent*; in particular, in any MPE σ,

$$\hat{V}_i^{\sigma,j}(s,T) = \begin{cases} \pi V_i^{\sigma,i}(s,T) + (1-\pi)V_i^{\sigma,-i}(s,T) & \text{if } j = i \\ (1-\pi)V_i^{\sigma,i}(s,T) + \pi V_i^{\sigma,-i}(s,T) & \text{if } j \neq i \end{cases}. \tag{3.6}$$

Combining (3.5) and (3.6), we have that, (ii), in the unique SPE σ for the game with $T = 1$ and for any $i, j \in \{1,2\}$,

$$\hat{V}_i^{\sigma,j}(s,1) = s_i + (1-\theta_s)\hat{W}_i^{\sigma,j}(1). \tag{3.7}$$

The analysis for the case $T = 1$ demonstrates salient features (i) and (ii) pointed out in the text. In fact, the analogues of these hold in general for $T \in \mathbb{N}$, as encapsulated below.

Proposition 12 (Finite Deadline Equilibrium). *Each of the following holds for any $T \in \mathbb{N}$:*

(i) There is a unique SPE σ which is stationary Markov perfect, with player strategies given by

$$A_1(T) = \left\{(\alpha,\beta) \in [0,1]^2 \mid \alpha \geq \underline{\alpha}^T := \delta\hat{W}_1^{\sigma,2}(T-1)\right\}$$
$$A_2(T) = \left\{(\alpha,\beta) \in [0,1]^2 \mid \alpha \leq \overline{\alpha}^T := 1 - \delta\hat{W}_2^{\sigma,1}(T-1)\right\},$$

and proposals in (α,β)-terms are independent of s and given by

$$x_1(T) = \left(\overline{\alpha}^T, [c_1']^{-1}\left(\frac{1-\delta\hat{W}^{\sigma,1}(T-1)}{1-\delta}\right)\right)$$
$$x_2(T) = \left(\underline{\alpha}^T, [c_2']^{-1}\left(\frac{1-\delta\hat{W}^{\sigma,2}(T-1)}{1-\delta}\right)\right).$$

(ii) For any $s \in X, i, j \in \{1, 2\}$,

$$\hat{V}_i^{\sigma,j}(s, T) = s_i + (1 - \theta_s)\hat{W}_i^{\sigma,j}(T). \tag{3.8}$$

Proposition 12 completely solves the finite deadline game. At any history, p always advances to her counterpart r a proposal which gives to r her reservation utiilty— $\delta\hat{W}_r^{\sigma,p}(T-1)$— from disagreeing. She does this by offering proportion $\delta\hat{W}_r^{\sigma,p}(T-1)$ of any *relative progress* $\beta_p(T)$ her proposal accomplishes. By advancing the project towards the frontier, agents reduce the additional surplus obtainable in the game after the current period and shape projects implemented in the future. In particular, relative progress β contracts the responder's *next-period prospects* $\hat{W}_r^{\sigma,p}(T-1)$ by factor $(1-\beta)$. Hence, the minimal amount r must be compensated in the current period is exactly $\beta \cdot \delta\hat{W}_r^{\sigma,p}(T-1)$, for any relative progress β the proposer advances. This gives the set of mutually acceptable policies

$$A^{\sigma}(T) := A_1^{\sigma,2}(T) \cap A_2^{\sigma,1}(T)$$

a conical shape.[4]

Moreover, the relative progress accomplished by the proposer in equilibrium is *decreasing* in the future total surplus $\delta\hat{W}^{\sigma,p}(T-1)$ in the game. It is useful to decompose this effect into two channels. First, as $\delta\hat{W}_r^{\sigma,p}(T-1)$ increases, we see from previous discussion that the leverage of the responder increases, which disincentivizes work in the current period, as a larger proportion of the gains from progress made by p must be ceded to r. Second, as $\hat{W}_p^{\sigma,p}$ increases, the incentives discussed in the previous paragraph also hold for the *proposer*; that is, as she advances the project forward she also forecloses her own future prospects by downsizing the stakes of the next period game. Hence, equilibrium effort $(\beta_p(T))$ is decreasing in the sum of these, $\delta\hat{W}^{\sigma,p}(T-1)$. This simple insight will also be key to the analysis of many of our main results.

[4]For instance, this feature of acceptance sets is also present in (Acharya and Ortner, 2021).

Next, by recursion, the strategies and equilibrium values of players delineated in Proposition 12 are fully determined. In particular, we have

$$W_i^{\sigma,i}(T) = \left(1 - \delta \hat{W}_{-i}^{\sigma,i}(T-1)\right) \beta_i(T) + \delta(1-\beta_i(T))\hat{W}_i^{\sigma,i}(T-1) - (1-\delta)c_i(\beta_i(T)) \quad (3.9)$$

$$W_{-i}^{\sigma,i}(T) = \delta \hat{W}_{-i}^{\sigma,i}(T-1)\beta_i(T) + \delta(1-\beta_i(T))\hat{W}_{-i}^{\sigma,i}(T-1) = \delta \hat{W}_{-i}^{\sigma,i}(T) \quad (3.10)$$

where we simply use (3.2), (3.8), and the transformation from Cartesian to (α, β)-coordinates. By simply summing (3.9) and (3.10), we obtain

$$W^{\sigma,i}(T) = \beta_i(T) + \delta(1-\beta_i(T))\hat{W}^{\sigma,i}(T-1) - (1-\delta)c_i(\beta_i(T)) \quad (3.11)$$

as the total surplus in the game when i is the proposer with deadline T.

Recursive Operators and Equilibrium Values

Here, we follow closely the treatment of value functions in (Acharya and Ortner, 2021) by expressing them in terms of recursive operators. Inspired by the form of strategies in Proposition 12 above, define

$$M_1(x) = [c_1']^{-1}\left(\frac{1-\delta x}{1-\delta}\right) + \delta\left(1 - [c_1']^{-1}\left(\frac{1-\delta x}{1-\delta}\right)\right)x - (1-\delta)c_1\left([c_1']^{-1}\left(\frac{1-\delta x}{1-\delta}\right)\right),$$
$$M_2(y) = [c_2']^{-1}\left(\frac{1-\delta y}{1-\delta}\right) + \delta\left(1 - [c_2']^{-1}\left(\frac{1-\delta y}{1-\delta}\right)\right)y - (1-\delta)c_2\left([c_2']^{-1}\left(\frac{1-\delta y}{1-\delta}\right)\right).$$

Then, let

$$\Phi(w,z) = (\varphi_1(w,z)\ ,\ \varphi_2(w,z)) := (\pi M_1(w) + (1-\pi)M_2(z)\ ,\ \pi M_2(z) + (1-\pi)M_1(w))\,. \quad (3.12)$$

And, furthermore,

$$\Psi(w,z) = (\psi_1(w,z)\ ,\ \psi_2(w,z)) := (M_1(\pi w + (1-\pi)z)\ ,\ M_2(\pi z + (1-\pi)w))\,. \quad (3.13)$$

Finally, let $\zeta : [0,1]^2 \times [0,1]^2 \to [0,1]^4$ be given by

$$\begin{aligned}\zeta(\mathbf{w},\mathbf{z}) &= \big(\zeta_1^1(\mathbf{w},\mathbf{z}), \zeta_2^1(\mathbf{w},\mathbf{z}), \zeta_1^2(\mathbf{w},\mathbf{z}), \zeta_2^2(\mathbf{w},\mathbf{z})\big) \\ &:= \big(M_1(\pi(w_1+w_2)+(1-\pi)(z_1+z_2)) - \delta(\pi w_2 + (1-\pi)z_2), \delta(\pi w_2 + (1-\pi)z_2), \\ &\quad \delta(\pi z_1 + (1-\pi)w_1), M_2(\pi(z_1+z_2)+(1-\pi)(w_1+w_2)) - \delta(\pi z_1 + (1-\pi)w_1)\big) \quad (3.14)\end{aligned}$$

The following is then an immediate corollary of Proposition 12—

Corollary 5. *1 For any deadline $T \in \mathbb{N}$, the unique SPE σ has*

(i) $(W^{\sigma,1}(T), W^{\sigma,2}(T)) = \Psi^T(0,0)$*;*

(ii) $(\hat{W}^{\sigma,1}(T), \hat{W}^{\sigma,2}(T)) = \Phi^T(0,0)$*; and,*

(iii) $(W_1^{\sigma,1}(T), W_2^{\sigma,1}(T), W_1^{\sigma,2}(T), W_2^{\sigma,2}(T)) = \zeta^T(\boldsymbol{0},\boldsymbol{0})$.

Corollary 5 is a helpful intermediate result. As we shall soon see, the technique of writing value functions in terms of recursive operators enables us to prove properties about equilibrium by simply analyzing features of these operators.

Limiting Equilibrium

The preceding discussion describes the unique subgame perfect equilibrium σ^T of each game with deadline T. We have seen that each such SPE is stationary Markov perfect. We now examine the limit as $T \to \infty$ of these equilibria. In particular, we show this limiting equilibrium is precisely the unique stationary Markov perfect equilibrium of the infinite deadline game.

Proposition 13 (Limiting Equilibrium).

(i) Any stationary MPE σ of the $T = \infty$-deadline game has equilibrium values that are the fixed points of Ψ, Φ, and ζ, respectively.

(ii) There is a unique stationary MPE σ^* of the $T = \infty$ game, which is the limit of the sequence $\{\sigma^T\}_{T=1}^{\infty}$ of the unique SPE of the finite deadline games.

Part (i) of the definition follows naturally from the definition of these operators and the recursivity inherent in stationary MPE. We prove part (ii) of the proposition by arguing that each of these operators is a contraction and using the contraction mapping theorem. Proposition 13 provides a foundation for focusing on the unique stationary MPE σ^* when considering equilibrium behavior in the infinite deadline game. σ^* approximates, for instance, behavior at σ^T when agents face T very large (a distant deadline). Because, in practice, all projects undertaken by agents have some practical stopping point, this provides a robustness to the main results presented in the next section. As alluded to in the previous subsection, we can use fixed point analysis on the operators Ψ, Φ, and ζ for comparative statics on elements of σ^*. We now turn to this endeavor.

3.4 Main Results

In this Section, I present the main result of the paper (Proposition 14), which describes how increasing the incumbency of a project leader affects her productivity. We show that the effect varies depending on the efficiency of the leader relative to her counterpart. Moreover, we relate this result to an important application: the relationship between government turnover and public spending.

3.4.1 Incumbency and Productivity

In this section, we examine the titular question of the paper, *how does the rate of incumbency π affect the productivity of any given project leader?* This question is relevant in a variety of settings. For instance, in organizational and work studies, there has been keen interest and debate on the effect job (in)security has on productivity. In the legislative context, this question has implications for how more frequent government turnover can affect progress on economic development, debt reduction, and other national goals or projects. Specifically, by focusing on a setting where agents have heterogenous costs and must cooperate to move the project forward, we also consider the possibility that incumbency incentives affect different *kinds* of project leaders differently.

Recall that in our model agent 1 is the efficient leader, having a cost which is a $\kappa < 1$ factor of her counterpart. In fact, this heterogeneity in costs has a crucial effect on how incumbency affects productivity for project leaders in equilibrium.

Proposition 14 (**Incumbency Effects**). *In the unique stationary MPE of the $T = \infty$ game, σ^*,*

(i) For any π, $W^{\sigma^*,1} > W^{\sigma^*,2}$

(ii) $W^{\sigma^*,1}$ and $\hat{W}^{\sigma^*,1}$ are both strictly increasing in π, while $W^{\sigma^*,2}$ and $\hat{W}^{\sigma^*,2}$ are both strictly decreasing in π.

(iii) The equilibrium effort exerted by agent 1 when she is the proposer, β_1^*, is strictly decreasing in π and that exerted by agent 2, β_2^*, is strictly increasing in π.

Each of (i) - (iii) provides a key description of equilibrium behavior in the infinite deadline game, and therefore, for games with sufficiently large (finite) deadlines T. Because 1 is the efficient project leader, it follows that surplus in the game is always larger when she is responsible for advancing the project, as given by (i). However, the intuition for (ii) and (iii)— why increased incumbency across both potential proposers yields *opposite* effects on their provision of effort— is non-trivial. It is also masked by the technical nature of the fixed point analysis used to prove these results in the Appendix. To elucidate, I discuss the driving forces behind (ii) and (iii) for equilibrium in the $T = 2$ case.[5]

When $T = 2$, only the next proposer receives any value from the game in the *next* period, and next period behavior is independent of π.[6] Increasing the rate of incumbency from π to $\pi' > \pi$ benefits the current proposer p through *two* channels: (1) first, it increases her next-period expected value by

$$\delta(\pi' - \pi)[W_p^{\sigma,p}(1) - W_p^{\sigma,r}(1)] = \delta(\pi' - \pi)W^{p,\sigma}(1).$$

(2) Simultaneously, it *decreases* the leverage of the respondent by lowering her next period expected value (her outside option from disagreeing) by $\delta(\pi' - \pi)W^{r,\sigma}(1)$. Note, however, that the proposer

[5] Proposition 14 does not imply the result for the $T = 2$ case, but each of (i) - (iii) does, in fact, hold, by inspection.
[6] See the base case of Proposition 12 proven in the text.

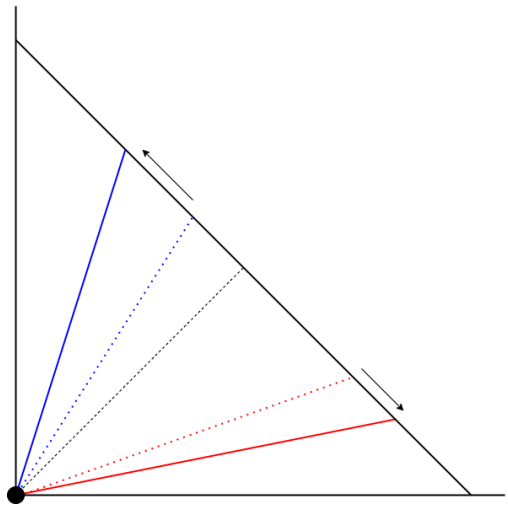


Figure 3.4.1: Here, $\overline{\alpha}^T$ is shown in blue and $\underline{\alpha}^T$ shown in red for two different values of π. The dotted lines represent these values for a smaller $\pi < \pi'$.

faces a *trade-off* from exerting effort; as greater effort is exerted, the total absolute size of the gains remaining to be had in the game shrink.

If, for p, the increase in leverage over the respondent is larger than the increase in the relative value of the next-period game, then she exerts greater effort in the current period as the comparative payoff per unit effort is higher. On the other hand, when the increase in relative value for the next-period game is greater, the proposer should reduce effort and increase the scale of the remaining final-period project. From the previous paragraph, the former is true for $p = 2$ and the latter for $p = 1$ since $W^{\sigma,1}(1) > W^{\sigma,2}(1)$.

The effect of π on the behavior or agents is a crucial factor in the design of *bargaining institutions*, since increasing the rate of incumbency not only alters the probability of a given sequence of proposers, but also changes the total contribution agents make to the project. In the following section, we provide a framework to study optimal choice over *ex ante* contractible incumbency rules.

3.4.2 Application: Government Turnover and Spending

We relate our work and main result to a simple model of government spending. Here, the project leader is an incumbent who requires the assent of an opposition (responder) in order to spend on

public goods. Ideological differences over which public goods are preferred (social safety net, military, etc.) lead agents to prefer different spending programs. We can think of the frontier in our model as a spending limit. Statutorily in the United States, for example, there is a cap on the amount of natinoal government debt that can be accrued— the *debt ceiling*— and many state provisions that similarly capt debt for local governments; this effectively limits the total budget of government. The cost of spending in our model can be thought of as a *political cost*— a reputational burden that a politician faces from increasing the polity's budget. Finally, note that by spending, agents *partially foreclose* the possibility of future spending by their successors. This effect is identified as a dynamic driving force of government debt in some of the existing theoretical literature — e.g., (Persson and Svensson, 1989), (Alesina and Tabellini, 1990), (Yared, 2019).

Proposition 14 then implies the following: if there is a party or politician for whom spending is *less costly*, when government turnover increases, she will spend *more*. On the other hand, the public servant for whom spending is more costly will spend less as the probability of turnover increases. There is some empirical evidence for this at at the local level of government. For instance, (Clinger et al., 2008) finds that turnover *increases* government debt in council–manager cities while it actually *decreases* government debt in council–mayor cities. (Stein, 1990) notes that in council–manager cities, managers' careers benefit from successful economic development, which often requires more government investment (and debt). It is hence plausible that the effective cost to a manager from implementing a spending plan of her own is less than that of the council, which faces electoral constraints; on the other hand, a mayor is even more publicly responsible for each dollar spent as the unique citywide politician. Hence, the empirical discrepancy (Clinger et al., 2008) finds is consistent with Proposition 14.

Finally, Proposition 14 also has implications for how overconfidence can affect spending. One can model *overconfidence* as an overestimation of skill— e.g., (Benoit and Dubra, 2011). In the context of our model, overconfident politicians may underestimate their reputational cost of public spending relative to their counterparts. As a result, overconfident politicians spend *more* when

government turnover increases, whereas underconfident politicians spend *less*. This interpretation rationalizes increases in spending from both parties.

3.5 Conclusion

In this article, I study a cooperative environment in which a project leader advances at personal cost a joint initiative forward with the assent of a collaborator. Agents have divergent preferences over the characteristics of the project, which allows for us to study bargaining interactions within a contributions game. I examine the role incumbency over project leadership has on leader's productivity. I find that incumbency enhances the productivity of the inefficient agent and limits that of the efficient contributor. I briefly relate these findings to literature on public spending.

The central finding of this paper has further ramifications for cooperative institutional settings. For instance, how do agents' preferences over incumbency affect the *structure* of bargaining institutions— that is, how often leadership rotates? It is therefore my hope that this article proves useful for further avenues of exploration.

Bibliography

Acharya, A. and Ortner, J. (2021). Paths to the frontier. *American Economic Journal: Microeconomics (forthcoming)*.

Admati, A. R. and Perry, M. (1991). Joint projects without commitment. *The Review of Economic Studies*, 58(2):259–276.

Alesina, A. and Tabellini, G. (1990). A positive theory of fiscal deficits and government debt. *The Review of Economic Studies*, 57(3):403–414.

Baron, D. P. (1996). A dynamic theory of collective goods programs. *The American Political Science Review*, 90(2):316–330.

Battaglini, M. and Coate, S. (2008). A dynamic theory of public spending, taxation, and debt. *American Economic Review*, 98(1):201–36.

Benoit, J.-P. and Dubra, J. (2011). Apparent overconfidence. *Econometrica*, 79(5):1591–1625.

Bonatti, A. and Rantakari, H. (2016). The politics of compromise. *American Economic Review*, 106(2):229–59.

Bowen, T. R., Chen, Y., and Eraslan, H. (2014). Mandatory versus discretionary spending: The status quo effect. *American Economic Review*, 104(10):2941–74.

Clinger, J. C., Feiock, R. C., McCabe, B. C., and Park, H.-J. (2008). Turnover, transaction costs, and time horizons: An examination of municipal debt financing. *The American Review of Public Administration*, 38(2):167–179.

Compte, O. and Jehiel, P. (2004). Gradualism in bargaining and contribution games. *The Review of Economic Studies*, 71(4):975–1000.

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

Dziuda, W. and Loeper, A. (2016). Dynamic collective choice with endogenous status quo. *Journal of Political Economy*, 124(4):1148–1186.

Eraslan, H., Evdokimov, K. S., and Zapal, J. (2020). Dynamic legislative bargaining. *ISER Discussion Paper*, No. 1090.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169–182.

Iaryczower, M., Oliveros, S., and Parihar, P. (2021). Endogenous contribution cycles. *Working Paper*.

Kessing, S. G. (2007). Strategic complementarity in the dynamic private provision of a discrete public good. *Journal of Public Economic Theory*, 9(4):699–710.

Kuvalekar, A. and Lipnowski, E. (2020). Job insecurity. *American Economic Journal: Microeconomics*, 12(2):188–229.

Persson, T. and Svensson, L. E. O. (1989). Why a stubborn conservative would run a deficit: Policy with time- inconsistent preferences. *The Quarterly Journal of Economics*, 104(2):325–345.

Stein, R. M. (1990). *Urban alternatives: Public and private markets in the provision of local services*. University of Pittsburgh.

Yared, P. (2019). Rising government debt: Causes and solutions for a decades-old trend. *Journal of Economic Perspectives*, 33(2):115–40.

3.A Omitted Proofs for Section 3

Proof of Proposition 12. We proceed by induction on T. Note that the base case $T = 1$ has already been proven in the text. Consider $T \geq 2$ then and suppose the statement holds for $T - 1$. Given a history h, let $h' = h \cup \{(-i, x, 1)\}$ and $h'' = h \cup \{(-i, x, 0)\}$. We first direct our attention to the acceptability sets, $A_i^\sigma(h, T)$ for a given player i. The necessary and sufficient condition for $x \in A_i^\sigma(h, T)$ is

$$\begin{aligned}
(1-\delta)x_i + \delta V_i^\sigma(h', T-1) &\geq (1-\delta)s_i + \delta V_i^\sigma(h'', T-1) \\
x_i + \delta(1-\theta_x)\hat{W}_i^{\sigma,-i}(T-1) &\geq s_i + \delta(1-\theta_s)\hat{W}_i^{\sigma,-i}(T-1), \text{ from inductive hypothesis} \\
\implies x_i - s_i &\geq \delta[(1-\theta_s) - (1-\theta_x)]\hat{W}_i^{\sigma,-i}(T-1) \\
\frac{x_i - s_i}{\theta_{x-s}} &\geq \delta\hat{W}_i^{\sigma,-i}(T-1). \qquad (3.15)
\end{aligned}$$

The set in $X(s)$ characterized by (3.15) is exactly $\{x \in X(s) \,|\alpha_{x-s} \geq \underline{\alpha}^T\}$ for $i = 1$ and $\{x \in X(s) \,|\alpha_{x-s} \leq \overline{\alpha}^T\}$ for $i = 2$. The maximization problem of the proposer (without loss of generality, 1) is thus given by

$$\begin{aligned}
&\max_{\alpha \geq \underline{\alpha}^T,\, \beta} \left\{ (1-\delta)\left(s_1 + (1-\theta_s)\left[\beta\alpha - c_1(\beta)\right]\right) + \delta\hat{V}_1^{\sigma,1}\left(s + (1-\theta_s)\left(\alpha\beta, (1-\alpha)\beta\right), T-1\right) \right\} \\
\simeq &\max_{\alpha \geq \underline{\alpha}^T,\, \beta} \Bigg\{ (1-\delta)\left(s_1 + (1-\theta_s)\left[\beta\alpha - c_1(\beta)\right]\right) + \\
&\qquad \delta\left(s_1 + (1-\theta_s)\beta\alpha + (1-\theta_s)(1-\beta)\hat{W}_1^{\sigma,1}(T-1)\right) \Bigg\}, \text{ using (iii) of inductive hyp.} \\
\simeq &\max_{\beta} \left\{ s_1 + (1-\theta_s)\beta\overline{\alpha}^T - (1-\delta)(1-\theta_s)c_1(\beta) + \delta(1-\theta_s)(1-\beta)\hat{W}_1^{\sigma,1}(T-1) \right\}, \\
&\text{by combining terms and noting that the objective is decreasing in } \alpha. \text{ By taking pos. affine trans.,} \\
\simeq &\max_{\beta} \left\{ \beta\overline{\alpha}^T - (1-\delta)c_1(\beta) + \delta(1-\beta)\hat{W}_1^{\sigma,1}(T-1) \right\} \qquad (3.16) \\
= &\max_{\beta} \left\{ \beta \cdot (1 - \delta\hat{W}_2^{\sigma,1}(T-1)) - (1-\delta)c_1(\beta) + \delta(1-\beta)\hat{W}_1^{\sigma,1}(T-1) \right\}. \qquad (3.17)
\end{aligned}$$

Using the first order conditions and noting the properties of c_1, maximization problem (3.16) has solution given by

$$\beta_1(h, T) = [c_1']^{-1}\left(\frac{1 - \delta\hat{W}^{\sigma,1}(T-1)}{1-\delta}\right). \qquad (3.18)$$

Note that because c_1 is strictly convex, and its argument in (3.18) is greater than 1, (3.18) has a unique, well-defined solution. Because the same procedure can be analogously applied for $i = 2$, we have shown the strategies in part (i) constitute the unique SPE, given the inductive hypothesis. Since the strategies are *not* a function of h, I drop this for ease of notation when referring to them. Now, as before,

$$(V_1^1(h,T), V_2^1(h,T)) = \Big(s_1 + (1-\theta_s)\left[\beta_1(T)(1 - \delta\hat{W}_2^{\sigma,1}(T-1)) - (1-\delta)c_1(\beta_1(T)) + \delta(1-\beta_1(T))\hat{W}_1^{\sigma,1}(T-1)\right],$$

$$s_2 + (1-\theta_s)\left[\delta\hat{W}_2^{\sigma,1}(T-1)\beta_1(T) + \delta(1-\beta_1(T))\hat{W}_2^{\sigma,1}(T-1)\right]\Big). \quad (3.19)$$

By substituting $h = h_1$ into (3.19),

$$(W_1^1(T), W_2^1(T)) = \Big(\beta_1(T)(1-\delta\hat{W}_2^{\sigma,1}(T-1)) - (1-\delta)c_1(\beta_1(T)) + \delta(1-\beta_1(T))\hat{W}_1^{\sigma,1}(T-1),$$
$$\beta_1(T)\cdot\delta\hat{W}_2^{\sigma,1}(T-1) + \delta(1-\beta_1(T))\hat{W}_2^{\sigma,1}(T-1)\Big). \quad (3.20)$$

And by substituting (3.20) back into (3.19), we obtain

$$V_i^1(s,T) = s_i + (1-\theta_s)W_i^1(T-1).$$

By analogy, again, the above also holds in the case of 2 the proposer. By using (3.6), we obtain (3.8), giving us statement (ii) of the Lemma, as desired. □

Proof of Corollary 5. We proceed by induction on T. In the base case $T = 1$, from the form of strategies in Proposition 12, $M_i(0) = \beta_i(1) + \delta(1-\beta_i(1))\hat{W}^{\sigma,i}(0) - (1-\delta)c_i(\beta_i(1))$, by the fact that $\hat{W}^{\sigma,i}(0) = 0$. Using (3.11),

$$\Psi(0,0) = (M_1(0), M_2(0)) = (W^{\sigma,1}(1), W^{\sigma,2}(1).$$

For (ii), we use the above to see directly

$$\Phi(0,0) = (\pi M_1(0) + (1-\pi)M_2(0), (1-\pi)M_1(0) + \pi M_2(0)) = (\hat{W}^{1,\sigma}(1), \hat{W}^{\sigma,2}(1)).$$

Finally, for (iii), we see that

$$\zeta(\mathbf{0},\mathbf{0}) = (M_1(0), 0, 0, M_2(0)) = ((W_1^{\sigma,1}(1), W_2^{\sigma,1}(1), W_1^{\sigma,2}(1), W_2^{\sigma,2}(1)).$$

The final equality results from the fact that $M_i(0) = W^{\sigma,i}(1) = W_i^{\sigma,i}(1)$, since as shown in the analysis for the $T = 1$ case in the text, the responder receives no surplus in this particular game. This completes the proof of the base case.

Next, suppose the statement holds for arbitrary $T - 1 \geq 1$. We prove the statement for T. By inductive hypothesis,

$$\begin{aligned}
\Psi^T(0,0) &= \Psi((\Psi^{T-1}(0,0))_1, (\Psi^{T-1}(0,0))_2) = \Psi(W^{\sigma,1}(T-1), W^{\sigma,2}(T-1)) \\
&= (M_1(\pi W^{\sigma,1}(T-1) + (1-\pi)W^{\sigma,2}(T-1)), M_2(\pi W^{\sigma,2}(T-1) + (1-\pi)W^{\sigma,1}(T-1))) \\
&= (\beta_1(T) + \delta(1-\beta_1(T))\hat{W}^{\sigma,1}(T-1) - (1-\delta)c_1(\beta_1(T)), \\
&\qquad \beta_2(T) + \delta(1-\beta_2(T))\hat{W}^{\sigma,2}(T-1) - (1-\delta)c_2(\beta_2(T)) \\
&= (W^{\sigma,1}(T), W^{\sigma,2}(T)),
\end{aligned}$$

as desired. On the other hand,

$$\Phi^T(0,0) = \Phi((\Phi^{T-1}(0,0))_1, (\Phi^{T-1}(0,0))_2) = \Phi(\hat{W}^{\sigma,1}(T-1), \hat{W}^{\sigma,2}(T-1))$$

$$
\begin{aligned}
&= (\pi M_1(\hat{W}^{\sigma,1}(T-1)) + (1-\pi)M_2(\hat{W}^{\sigma,2}(T-1)),\\
&\qquad \pi M_2(\hat{W}^{\sigma,2}(T-1)) + (1-\pi)M_1(\hat{W}^{\sigma,1}(T-1))\\
&= (\hat{W}^{\sigma,1}(T), \hat{W}^{\sigma,2}(T)).
\end{aligned}
$$

Finally, using the previous results,

$$
\begin{aligned}
\zeta^T(\mathbf{0},\mathbf{0}) &= \zeta(\zeta^{T-1}(\mathbf{0},\mathbf{0})) = \zeta\left(W_1^{\sigma,1}(T-1), W_2^{\sigma,1}(T), W_1^{\sigma,2}(T-1), W_2^{\sigma,2}(T-1)\right)\\
&= \Bigg(M_1(\hat{W}^{\sigma,1}(T-1)) - \delta\hat{W}_2^{\sigma,1}(T-1), \delta\hat{W}_2^{\sigma,1}(T-1),\\
&\qquad \delta\hat{W}_1^{\sigma,2}(T-1), M_2(\hat{W}^{\sigma,2}(T-1)) - \delta\hat{W}_1^{\sigma,2}(T-1)\Bigg)\\
&= \left(W_1^{\sigma,1}(T-1), W_2^{\sigma,1}(T-1), W_1^{\sigma,2}(T-1), W_2^{\sigma,2}(T-1)\right),
\end{aligned}
$$

where the final equality results from (3.10) and the definition of

$$
W^{\sigma,i}(T-1) = M_i(\hat{W}^{\sigma,i}(T-1)).
$$

□

Proof of Proposition 13. (i) Consider any stationary MPE σ of the infinite deadline game. Fix 1 as the proposer and 2 the responder, without loss of generality. From Observation 1— which follows directly from the definition— strategies of players under σ history-invariant in (α,β)-space. Hence, $x_1(\infty) = (\alpha_p(\infty), \beta_p(\infty))$ satisfies

$$
(\alpha_1(\infty), \beta_1(\infty)) \in \text{argmax}_{(\alpha_1,\beta_1)\in A_2^\sigma(\infty)} \alpha_1\cdot\beta_1 + \delta(1-\beta_1)\hat{W}_1^{\sigma,1}(\infty) - (1-\delta)c_1(\beta_1). \quad (3.21)
$$

And $A_2^\sigma(\infty)$ satsifies

$$
\begin{aligned}
(\alpha_1,\beta_1) \in A_2^\sigma(\infty) &\Leftrightarrow (1-\alpha_1)\beta_1 + (1-\beta_1)\delta\hat{W}_2^{\sigma,1}(\infty) \geq \delta\hat{W}_2^{\sigma,1}(\infty)\\
&\Leftrightarrow \alpha_1 \leq 1 - \delta\hat{W}_2^{\sigma,1}(\infty)
\end{aligned}
$$

By the fact that (3.21) is strictly increasing in α_1, it follows that $\alpha_1(\infty) = 1 - \delta\hat{W}_1^{\sigma,1}$, and the maximization problem of the proposer, agent 1, can be written

$$
\beta_1(\infty) \in \text{argmax}_{\beta_1\in[0,1]}\left\{(1-\delta\hat{W}_2^{\sigma,1}(\infty))\beta_1 + \delta(1-\beta_1)\hat{W}_1^{\sigma,1}(\infty) - (1-\delta)c_1(\beta_1)\right\},
$$

which has solution

$$
\beta_1(\infty) = [c_1']^{-1}\left(\frac{1-\delta\hat{W}^{\sigma,1}(\infty)}{1-\delta}\right), \quad (3.21)
$$

as a function of the equilibrium continuation value. The case of agent 2 as the proposer is, of course, exactly analogous. Consider the continuation values associated with this stationary

MPE σ.From above, for each $i \in \{1,2\}$,

$$\begin{aligned}
W^{\sigma,i}(\infty) &= [c_i']^{-1}\left(\frac{1-\delta(\pi W^{\sigma,i}(\infty)+(1-\pi)W^{\sigma,-i}(\infty))}{1-\delta}\right) - (1-\delta)c_1\left(\delta(\pi W^{\sigma,i}(\infty)+(1-\pi)W^{\sigma,-i}(\infty))\right) \\
&\quad + \left(1-[c_1']^{-1}\left(\frac{1-\delta(\pi W^{\sigma,i}(\infty)+(1-\pi)W^{\sigma,-i}(\infty))}{1-\delta}\right)\right)\delta\left(\pi W^{\sigma,i}(\infty)+(1-\pi)W^{\sigma,-i}(\infty)\right) \\
&= \psi_i(W^{\sigma,i}(\infty), W^{\sigma,-i}(\infty)).
\end{aligned}$$

This of course gives us $(W^{\sigma,1}(\infty), W^{\sigma,2}(\infty)) = \Psi\left(W^{\sigma,1}(\infty), W^{\sigma,2}(\infty)\right)$— the desired fixed point condition for Ψ. Furthermore,

$$\hat{W}^{\sigma,i}(\infty) = \pi M_i(\hat{W}^{\sigma,i}(\infty)) + (1-\pi)M_{-i}(\hat{W}^{\sigma,-i}(\infty)) = \varphi_i(\hat{W}^{\sigma,i}(\infty), \hat{W}^{\sigma,-i}(\infty)).$$

Hence, $(\hat{W}^{\sigma,1}(\infty), \hat{W}^{\sigma,2}(\infty))$ is a fixed point of Φ. Finally,

$$\begin{aligned}
W_i^{\sigma,i}(\infty) &= (1-\delta\hat{W}_{-i}^{i}(\infty))\cdot[c_i']^{-1}\left(\frac{1-\delta\hat{W}^{\sigma,i}(\infty)}{1-\delta}\right) - (1-\delta)c_i\left([c_i']^{-1}\left(\frac{1-\delta\hat{W}^{\sigma,i}(\infty)}{1-\delta}\right)\right) + \\
&\qquad \delta\left(1-[c_i']^{-1}\left(\frac{1-\delta\hat{W}^{\sigma,i}(\infty)}{1-\delta}\right)\right)\hat{W}_i^{\sigma,i}(\infty) \\
&= M_i(\hat{W}^{\sigma,i}(\infty)) - \delta\hat{W}_{-i}^{\sigma,i}(\infty), \qquad \text{and} \\
W_{-i}^{\sigma,i}(\infty) &= \delta\hat{W}_{-i}^{\sigma,i}(\infty)\cdot[c_i']^{-1}\left(\frac{1-\delta\hat{W}^{\sigma,i}(\infty)}{1-\delta}\right) + \left(1-[c_i']^{-1}\left(\frac{1-\delta\hat{W}^{\sigma,i}(\infty)}{1-\delta}\right)\right)\delta\hat{W}_{-i}^{\sigma,i}(\infty) \\
&= \delta\hat{W}_{-i}^{\sigma,i}(\infty).
\end{aligned}$$

Therefore, $(W_i^{\sigma,j}(\infty))_{i,j\in\{1,2\}^2}$ is a fixed point of ζ, completing the proof of (i).

(ii) We show that each of the operators Ψ, Φ, and ζ is a contraction.

$$\begin{aligned}
\frac{\partial}{\partial w}\psi_1(w,z) &= \pi M_1'(\pi w+(1-\pi)z) \\
&= \pi\left[\frac{1}{(1-\delta)c_1''\left(\frac{1-\delta(\pi w+(1-\pi)z)}{1-\delta}\right)}\left(-\delta+\delta^2(\pi x+(1-\pi)y)+\delta(1-\delta(\pi w+(1-\pi)z))\right)\right. \\
&\qquad \left. + \delta\left(1-[c_1']^{-1}\left(\frac{1-\delta(\pi w+(1-\pi)z)}{1-\delta}\right)\right)\right] \\
&= \delta\pi\left(1-[c_1']^{-1}\left(\frac{1-\delta(\pi w+(1-\pi)z)}{1-\delta}\right)\right) \\
\frac{\partial}{\partial z}\psi_1(w,z) &= (1-\pi)M_1'(\pi w+(1-\pi)z) = \delta(1-\pi)\left(1-[c_1']^{-1}\left(\frac{1-\delta(\pi w+(1-\pi)z)}{1-\delta}\right)\right).
\end{aligned}$$

By analogous calculation,

$$\begin{aligned}
\frac{\partial}{\partial w}\psi_2(w,z) &= \delta(1-\pi)\left(1-[c_2']^{-1}\left(\frac{1-\delta(\pi z+(1-\pi)w)}{1-\delta}\right)\right) \\
\frac{\partial}{\partial z}\psi_2(w,z) &= \delta\pi\left(1-[c_2']^{-1}\left(\frac{1-\delta(\pi z+(1-\pi)w)}{1-\delta}\right)\right)
\end{aligned}$$

It is a well-known fact that the Lipschitz constant of Ψ is then

$$\sup_{\mathbf{y}\in[0,1]^2} ||J_\Psi(\mathbf{y})||_o,$$

where J_Ψ is the Jacobian of Ψ and $||\cdot||_o$ is the standard operator norm. From the derivatives above, it is clear that $||J_\Psi(\mathbf{y})||_o \leq \delta \cdot \max\{\pi, 1-\pi\}$. Hence,

$$|\Psi(\mathbf{y}) - \Psi(\mathbf{y'})| \leq \delta \cdot \max\{\pi, 1-\pi\}|\mathbf{y} - \mathbf{y'}|$$

so that Ψ is a contraction mapping. The arguments for the other two operators Φ and ζ are analogous, directly calculating the derivative with respect to each argument and showing that any partial derivative is always less strictly than 1.

Since each of Ψ, Φ, and ζ is a contraction, each has a unique fixed point. But now, from (i), the strategies (i) and (i), defined consistent with any stationary MPE, are uniquely determined. We therefore have a unique stationary MPE σ^* with continuation values given by the *unique* fixed points of Ψ, Φ, and ζ.

To conclude, by the Contraction Mapping (Banach Fixed Point) Theorem, the sequences $\Psi^T(0,0), \Phi^T(0,0)$, and $\zeta^T(\mathbf{0},\mathbf{0})$ all converge to the continuation values $(W^{\sigma^*,1}(\infty), W^{\sigma^*,2}(\infty))$, $(\hat{W}^{\sigma^*,1}(\infty), \hat{W}^{\sigma^*,2}(\infty))$, and $(W_i^{\sigma^*,j})_{(i,j)\in\{1,2\}^2}$, respectively, corresponding to σ^*. Hence, σ^* is not only the unique stationary MPE, it is also the limiting equilibrium of the sequence $\{\sigma^T\}_{T=1}^{\infty}$of unique SPE of each finite deadline game.

□

3.B Omitted Proofs for Section 4

Proof of Proposition 14. For the proof, we write $W^{\sigma^*,i}(\pi)$, $\hat{W}^{\sigma^*,i}(\pi)$, etc., for the equilibrium continuation values in the game to underline their dependence on π.

(i) Consider first the case $\pi \geq \frac{1}{2}$. Recall that $\left(W^{\sigma^*,1}, W^{\sigma^*,2}\right) = \lim_{T\to\infty} \Psi^T(0,0)$. We show that if $\pi \geq \frac{1}{2}$, then $(\Psi^T(0,0))_1 > (\Psi^T(0,0))_2$, for any $T \in \mathbb{N}$. Directly, $\psi_1(0,0) = M_1(0) > M_2(0) = \psi_2(0,0)$, so that this is true for $T = 1$. Next, suppose the statement holds for arbitrary $T - 1 \geq 1$. Since $\pi \geq \frac{1}{2}$, by inductive hypothesis,

$$\pi(\Psi^{T-1}(0,0))_1 + (1-\pi)(\Psi^{T-1}(0,0))_2 > \pi(\Psi^{T-1}(0,0))_2 + (1-\pi)(\Psi^{T-1}(0,0))_1.$$

Hence, because M_i are increasing functions,

$$\begin{aligned}(\Psi^T(0,0))_1 &= M_1(\pi(\Psi^{T-1}(0,0))_1 + (1-\pi)(\Psi^{T-1}(0,0))_2) \\ &> M_1(\pi(\Psi^{T-1}(0,0))_2 + (1-\pi)(\Psi^{T-1}(0,0))_1) \\ &> M_2(\pi(\Psi^{T-1}(0,0))_2 + (1-\pi)(\Psi^{T-1}(0,0))_1) = (\Psi^T(0,0))_2.\end{aligned}$$

Therefore, $(\Psi^T(0,0))_1 > (\Psi^T(0,0))_2$, for every $T \in \mathbb{N}$ and so for their limit, $W^{\sigma^*,1} \geq W^{\sigma^*,2}$. However, if $W^{\sigma^*,1} = W^{\sigma^*,2} = w^*$, then $M_1(w^*) > M_2(w^*)$ and such a pair (w^*, w^*) cannot be a fixed point of Ψ. We hence have the result $W^{\sigma^*,1} > W^{\sigma^*,2}$ for $\pi \geq \frac{1}{2}$.

Next, consider $\pi < \frac{1}{2}$, and suppose for the sake of contradiction, that $W^{\sigma^*,1} \leq W^{\sigma^*,2}$. So,

$$\pi W^{\sigma^*,1} + (1-\pi)W^{\sigma^*,2} \geq \pi W^{\sigma^*,2} + (1-\pi)W^{\sigma^*,1}.$$

But, then,

$$\psi_1(W^{\sigma^*,1}, W^{\sigma^*,2}) = M_1(\pi W^{\sigma^*,1} + (1-\pi)W^{\sigma^*,2}) > M_2(\pi W^{\sigma^*,2} + (1-\pi)W^{\sigma^*,1}) = \psi_2(W^{\sigma^*,1}, W^{\sigma^*,2}).$$

And because $\left(W^{\sigma^*,1}, W^{\sigma^*,2}\right)$ is a fixed point of Ψ, this generates a contradiction. Therefore, we must have $W^{\sigma^*,1} > W^{\sigma^*,2}$, completing the proof of (i).

(ii) From Proposition 13,

$$\hat{W}^{\sigma^*,1}(\pi) = \varphi_1(\hat{W}^{\sigma^*,1}(\pi), \hat{W}^{\sigma^*,2}(\pi)) \quad \text{and} \quad \hat{W}^{\sigma^*,2}(\pi) = \varphi_2(\hat{W}^{\sigma^*,1}(\pi), \hat{W}^{\sigma^*,2}(\pi)).$$

Using the Implicit function theorem and the definition of φ_i,

$$\begin{aligned}\frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,1}(\pi) = M_1(\hat{W}^{\sigma^*,1}(\pi)) + \pi\frac{\partial}{\partial w}M_1|_{w=\hat{W}^{\sigma^*,1}(\pi)} \cdot \frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,1}(\pi) - M_2(\hat{W}^{\sigma^*,2}(\pi)) \\ + (1-\pi)\frac{\partial}{\partial z}M_2|_{z=\hat{W}^{\sigma^*,2}(\pi)} \cdot \frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,2}(\pi), \text{ and,} \qquad (3.22)\\ \frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,2}(\pi) = M_2(\hat{W}^{\sigma^*,2}(\pi)) + \pi\frac{\partial}{\partial z}M_2|_{z=\hat{W}^{\sigma^*,2}(\pi)} \cdot \frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,2}(\pi) - M_1(\hat{W}^{\sigma^*,1}(\pi)) \\ + (1-\pi)\frac{\partial}{\partial w}M_1|_{w=\hat{W}^{\sigma^*,1}(\pi)} \cdot \frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,1}(\pi). \qquad (3.23)\end{aligned}$$

We can view (3.22), (3.23) as a system of equations for $\left(\frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,1}(\pi), \frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,2}(\pi)\right)$, in which case substitution gives us

$$\begin{aligned}&\frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,1}(\pi) = M_1(\hat{W}^{\sigma^*,1}(\pi)) - M_2(\hat{W}^{\sigma^*,2}(\pi)) + \pi\frac{\partial}{\partial w}M_1|_{w=\hat{W}^{\sigma^*,1}(\pi)} \cdot \frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,1}(\pi) + \\ &\frac{(1-\pi)\frac{\partial}{\partial z}M_2|_{z=\hat{W}^{\sigma^*,2}(\pi)}}{1-\pi\frac{\partial}{\partial z}M_2|_{z=\hat{W}^{\sigma^*,2}(\pi)}}\left[M_2(\hat{W}^{\sigma^*,2}(\pi)) - M_1(\hat{W}^{\sigma^*,1}(\pi)) + (1-\pi)\frac{\partial}{\partial w}M_1|_{w=\hat{W}^{\sigma^*,1}(\pi)} \cdot \frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,1}(\pi)\right]\end{aligned}$$

Rearranging to solve for $\frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,1}(\pi)$, we obtain

$$\frac{\partial}{\partial\pi}\hat{W}^{\sigma^*,1}(\pi) = \frac{\left[M_1(\hat{W}^{\sigma^*,1}(\pi)) - M_2(\hat{W}^{\sigma^*,2}(\pi))\right]\left(\frac{1-\frac{\partial}{\partial z}M_2|_{z=\hat{W}^{\sigma^*,2}}}{1-\pi\frac{\partial}{\partial z}M_2|_{z=\hat{W}^{\sigma^*,2}}}\right)}{1-\frac{\partial}{\partial w}M_1|_{w=\hat{W}^{\sigma^*,1}(\pi)}\left(\pi + (1-\pi)\frac{1-\frac{\partial}{\partial z}M_2|_{z=\hat{W}^{\sigma^*,2}}}{1-\pi\frac{\partial}{\partial z}M_2|_{z=\hat{W}^{\sigma^*,2}}}\right)}. \qquad (3.24)$$

From the proof of Proposition 13, recall that

$$\frac{\partial}{\partial w}M_i(w) = \delta[c_i']^{-1}\left(\frac{1-\delta w}{1-\delta}\right) < 1.$$

It therefore follows that

$$0 < \frac{1 - \frac{\partial}{\partial z} M_2|_{z=\hat{W}^{\sigma^*,2}}}{1 - \pi \frac{\partial}{\partial z} M_2|_{z=\hat{W}^{\sigma^*,2}}} < 1,$$

and furthermore the denominator of (3.24) is therefore strictly positive. In fact, from (3.24), it must be that

$$\frac{\partial}{\partial \pi} \hat{W}^{\sigma^*,1}(\pi) > 0 \Leftrightarrow M_1(\hat{W}^{\sigma^*,1}(\pi)) - M_2(\hat{W}^{\sigma^*,2}(\pi)) > 0.$$

Of course, from the proof of (i) above, we have $M_1(\hat{W}^{\sigma^*,1}(\pi)) = W^{\sigma^*,1}(\pi) > W^{\sigma^*,2}(\pi) = M_2(\hat{W}^{\sigma^*,2}(\pi)$, for any π. Therefore, $\frac{\partial}{\partial \pi} \hat{W}^{\sigma^*,1}(\pi) > 0$.

Finally, summing (3.22) and (3.23) immediately gives us

$$\left(1 - \frac{\partial}{\partial w} M_1|_{w=\hat{W}^{\sigma^*,1}(\pi)}\right) \frac{\partial}{\partial \pi} \hat{W}^{\sigma^*,1}(\pi) = -\left(1 - \frac{\partial}{\partial z} M_2|_{z=\hat{W}^{\sigma^*,2}}\right) \frac{\partial}{\partial \pi} \hat{W}^{\sigma^*,2}(\pi)$$
$$\implies \text{sgn}\left(\frac{\partial}{\partial \pi} \hat{W}^{\sigma^*,1}(\pi)\right) = -\text{sgn}\left(\frac{\partial}{\partial \pi} \hat{W}^{\sigma^*,2}(\pi)\right),$$

so that $\frac{\partial}{\partial \pi} \hat{W}^{\sigma^*,2}(\pi) < 0$, completing the proof of (ii).

(iii) This follows directly from the result of (ii). Note that c_i' is strictly increasing— i.e., c_i is convex— and

$$\beta_i^* = [c_i']^{-1}\left(\frac{1 - \delta \hat{W}^{\sigma^*,i}(\pi)}{1 - \delta}\right),$$

which implies the result.

□