Joint Discussion Paper Series in Economics

by the Universities of
Aachen • Gießen • Göttingen
Kassel • Marburg • Siegen

ISSN 1867-3678

No. 11-2026

Fabian Mankat and Konstantinos Theocharopoulos

A theory of the adverse effects of repression

Coordination: Bernd Hayo
Philipps-University Marburg
School of Business and Economics
Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

A theory of the adverse effects of repression*

Fabian Mankat†[1] and Konstantinos Theocharopoulos‡[2]

[1]University of Kassel

[2]University of Siegen

August 2026

Abstract

Incumbents with the capacity to repress often stop well short of using it fully. We address this puzzle with a dynamic model of opposition behavior in regimes where contested elections coexist with systematic advantages for those in power. Opposition factions choose whether to become politically active under repression and, once active, whether to cooperate with the other active factions or pursue independent paths. Repression deters participation and shrinks the active opposition, but among the factions that remain it strengthens the incentive to cooperate, since victory ends repression and a smaller field makes a united front easier to sustain. At low and high levels of repression, entry incentives govern the size of the opposition and additional repression works as intended. At intermediate levels, the capacity to sustain cooperation governs it and additional repression enlarges the united front. When unity poses a sufficient threat and factions are sufficiently patient, the probability of staying in power is non-monotone in repression, and repressing as much as possible can be strictly suboptimal.

Keywords: Repression, hybrid regimes, opposition, incumbent survival

JEL: C73, D72, D74

*We thank Frank Bohn, Björn Frank, Carsten Hefeker, Otto Swank, Georg von Wangenheim, and Stephane Wolton for helpful comments and suggestions.

†University of Kassel, fabian.mankat@uni-kassel.de.

‡University of Siegen, konstantinos.theocharopoulos@uni-siegen.de.

1 Introduction

Repression is a principal instrument by which incumbents in nondemocratic settings preserve power. It encompasses state actions that raise the cost of political activity, or that restrict such activity directly, against actual or potential challengers (Davenport, 2007; Poe & Tate, 1994). The straightforward expectation is that incumbents with substantial repressive capacity have strong incentives to deploy it against opposition threats (Ritter & Conrad, 2016; Young, 2019). Many do not. They stop well short of the highest feasible level, combining selective or intermediate repression with other instruments of control while leaving a degree of open contestation intact (Bove et al., 2017; Gandhi, 2015; Guriev & Treisman, 2023).[1] The pattern poses a puzzle: if repression discourages political activity, why is maximal repression so often not the safest strategy for incumbents?

We address this question through a theoretical model of opposition behavior under repression. The novel insight is that an increase in repression may sometimes have an adverse effect on the incumbent's probability of staying in power, which can make the maximal level of repression suboptimal even when repression carries no external cost. The underlying reason is that repression does not only deter factions from becoming active, it also strengthens active factions' incentives to cooperate with one another, which may backfire on the incumbent.

The model formalizes this argument in a dynamic setting: repression is set at a fixed level, and potential opposition factions first choose whether to become politically active, a step that exposes them to repression in every subsequent period. Active factions then choose in every period whether to cooperate, by adopting a united oppositional stance, or to defect, by pursuing an independent path at the expense of the rest of the opposition. An electoral contest determines whether the incumbent stays in power.[2]

Evaluating the incumbent's prospects at the equilibrium least favorable

[1] The relationship between repression and dissent is empirically ambiguous. Lichbach (1995) and Moore (2000) discuss how repression can both deter and provoke opposition activity. Pierskalla (2010) models this non-monotonic pattern directly. Fein (1995) documents that partly free states exhibit worse life-integrity violations than fully authoritarian ones, suggesting a similar relationship between political openness and state violence.

[2] Although we model the contest as an election, the underlying logic extends to other forms of political contestation, including mass protests, riots, and armed challenges. We focus on the electoral case because it provides the most natural setting for the mechanisms we study.

to the incumbent, we establish the following two main results. First, the incumbent's probability of staying in power can be non-monotone in repression. At low levels of repression, participation is cheap, the opposition is large and fragmented, and additional repression works as intended: it deters entry and strengthens the incumbent. At high levels, the opposition presents a united front but participation is so costly that entry incentives again govern the size of the opposition. Additional repression deters entry and strengthens the incumbent once more. In between, the size of the opposition is governed not by how many factions wish to enter but by how many can sustain cooperation, and additional repression enlarges the united front that confronts the incumbent. Repression is counterproductive precisely at those intermediate levels, strong enough to have thinned the fragmented opposition, not strong enough to deter the united one. The described unifying pressure that repression exerts on the opposition is present for any parameter values, and translates into a non-monotone winning probability of the incumbent whenever the advantage against a fragmented opposition sufficiently exceeds the advantage against a united one, and factions are sufficiently patient. Second, this non-monotonicity carries a direct implication for the incumbent's choice: whenever the maximal feasible level of repression falls into some intermediate region, an incumbent maximizing the probability of staying in power represses strictly less than possible.

Our argument applies most directly to hybrid regimes, political systems that combine democratic and authoritarian elements, occupying the gray zone between the two. Formally democratic institutions, especially multiparty elections, coexist with systematic incumbent advantages and meaningful constraints on opposition competition.[3] In these settings, opposition factions can organize, campaign, and contest office, while the incumbent benefits from control of state resources and administrative power, with repression available on top of these (Howard & Roessler, 2006; Hyde & Marinov, 2012). Elections serve both as instruments of nondemocratic control and as potential focal points for opposition cooperation (Fearon, 2011; Gandhi & Przeworski, 2007).[4] Our analysis

[3]These regimes have been studied under various labels, including hybrid regime, semi-authoritarian regime, competitive authoritarian regime, and electoral autocracy. We use hybrid regime throughout. Gandhi and Lust-Okar (2009), Levitsky and Way (2010), Schedler (2013), and Diamond (2002) discuss the terminology and classification of these regimes and the role of elections within them.

[4]Wahman (2013) examines the electoral consequences of opposition coalitions in authoritarian settings, while Greene (2007), Magaloni (2006), and Reuter (2017) study hegemonic

isolates repression as the sole instrument that shapes opposition behavior in this environment. Because factions remain active in politics, the incumbent's problem extends beyond deterring participation. It includes managing the possibility that repression leaves a smaller but more cooperative opposition.

Existing approaches address parts of our research question, without fully resolving it. One strand emphasizes that repression can provoke resistance, distort the information available to rulers, or generate moral hazard problems in political control (Dragu & Przeworski, 2019; Gibilisco, 2021; Hollyer et al., 2015; Montagnes & Wolton, 2019; Rosenfeld & Wallace, 2024; Slantchev & Matush, 2020; Tyson, 2018; Wintrobe, 1990). Recent formal work traces such trade-offs dynamically, showing how repression that quells present threats comes at the expense of policies that would reduce future mobilization (Di Lonardo et al., 2026), and how the reach of repression can vary non-monotonically with features of the environment such as inequality (Roy & MacDonald, 2026). Another strand documents the diverse political responses repression elicits, from anger and renewed activism to obedience and even expressions of loyalty (Chiang, 2021; Hager & Krakowski, 2022; LeBas & Young, 2024; Mitchell et al., 2026; Robertson, 2011; Rozenas & Zhukov, 2019; Sullivan & Davenport, 2017). These contributions advance the study of nondemocratic politics broadly, examining its many institutional forms, but they typically treat the opposition in aggregate.

Hybrid regimes, however, combine the institutional features of nondemocratic rule with the multiparty competition characteristic of democracies. Opposition politics is therefore organized through parties, leaders, and factions whose choices are interdependent, and repression bears on two distinct problems these actors face. The first concerns participation: work on collective action under repression asks when individuals and groups become politically active despite the costs that repression imposes (E. Bueno de Mesquita & Shadmehr, 2023; Morris & Shadmehr, 2024; Shadmehr & Bernhardt, 2011; Siegel, 2011; Tucker, 2007). The second concerns the interaction of those who are active: studies of opposition fragmentation examine why factions fail to unite even when unity would serve them all (Arriola et al., 2021; Buckles, 2019; Hassan, 2024), and related work shows how regimes deepen these divisions deliberately, with targeted repression undermining joint action across groups and information

party rule and opposition fragmentation. Helms (2023) reviews opposition politics across democratic and authoritarian regimes.

manipulation impeding mass coordination for political change (Edmond, 2013; Rozenas, 2020).[5] These two problems have largely been studied on their own, and the question of how repression reshapes cooperation incentives among the factions that remain active is underexplored. Our distinctive step is to bring the two responses together in a single dynamic framework and trace their joint consequences for the incumbent: repression deters participation at the first decision and pushes those who remain active toward cooperation at the second. This two-part response gives rise to the ambiguous effects of repression for the incumbent, and it is what aggregate accounts of repression and dissent do not capture.

Our analysis also complements a growing literature on how incumbents combine calibrated repression with institutional design, coalition management, information control, and economic policy (Acemoglu & Robinson, 2005; Acemoglu et al., 2008, 2026; Boix & Svolik, 2013; B. Bueno de Mesquita et al., 2003; Lagunoff, 2025; Meng, 2021; Svolik, 2012; Wright & Escribà-Folch, 2012).[6] These studies show that incumbents often stay in power through the choices they make within a constrained political environment. We focus on one problem within that setting: how the opposition's twofold response to a given level of repression shapes the incumbent's prospects for staying in power.

The paper proceeds as follows. Section 2 motivates the analysis through illustrative cases. Section 3 introduces the model, which describes the interaction of opposition factions as a dynamic game, and Section 4 studies it. At these stages the incumbent is treated as a passive player who sets a fixed level of repression but takes no further action, so the analysis focuses on how the factions interact among themselves. Section 5 brings the incumbent back in from a mechanism design standpoint, analyzing how the incumbent's probability of staying in power varies with repression, given how the factions respond.[7] Section 6 discusses the results, and Section 7 concludes.

[5] Related arguments about information control and strategic communication in authoritarian settings appear in Gehlbach (2010), Guriev and Treisman (2019), King et al. (2013), Lorentzen (2014), Roberts (2018), and Xu (2021). Guriev and Treisman (2020) develops the logic of informational autocracy, showing that leaders substitute information manipulation for overt repression when the informed elite is of intermediate size. Egorov and Sonin (2024) and Gehlbach et al. (2016) survey formal approaches to nondemocratic politics, while Przeworski (2023) offers a critique of this modeling tradition.

[6] Escribà-Folch (2013) provides cross-national evidence linking repression strategies to autocratic survival, while Esberg (2021) analyzes anticipatory repression before threats fully materialize.

[7] Our formulation is equivalent to treating the incumbent as an active player who selects

2 Motivating cases

Before turning to the formal model, we examine cases that motivate the theoretical argument: Russia under Vladimir Putin, Turkey under Recep Tayyip Erdoğan, Venezuela under Hugo Chávez and later Nicolás Maduro, and Hungary under Viktor Orbán. Each qualifies as a hybrid regime during the period we consider (Bozóki & Hegedűs, 2018; Corrales, 2020; Esen & Gümüşçü, 2016; Gel'man, 2015; Levitsky & Way, 2010). These cases are illustrations rather than tests of the theory, and they do not exhaust the considerations that shape political conflict in such regimes. Each reflects forces well beyond the logic studied here, among them elite coalition management, immediate threat responses, ideological projects, international pressure, and economic conditions. Across the cases, however, the same pattern of opposition responses recurs: factions adjust their participation in response to the political environment they face, those that remain active choose how to interact with one another, and the resulting opposition shapes the incumbent's prospects for staying in power.

We present the cases in turn. Russia and Venezuela show the pattern at opposite ends of the range, the first under moderate repression with a fragmented opposition, the second under escalating repression that eventually drew the opposition together. Turkey is of particular interest because it provides within-country variation across the full range of repression. Hungary is presented last because the dominant instruments of incumbent consolidation differ from those in the other cases, while the opposition response itself remains comparable.

We begin with Russia. Between 2004 and 2011, Russia under Vladimir Putin shows a moderate-repression equilibrium: a fragmented opposition alongside a secure incumbent. Repression was real but contained, working through selective prosecutions, restrictions on independent media, and pressure on those who funded the opposition, narrowing the space for competition without closing it (Gel'man, 2015; Treisman, 2011). Several parties remained active in national politics, the Communist Party of the Russian Federation and the Liberal Democratic Party of Russia inside the State Duma, Yabloko and the Union of Right Forces outside it after losing parliamentary representation in

the level of repression once at the start of the game. Since Section 5 characterizes the incumbent's probability of staying in power at every level of repression, it delivers precisely the objective such an incumbent would maximize. We keep repression exogenous in the model itself only to spare the notation an additional player and stage.

2003. Each held its own corner of the political spectrum and none built a common platform against the incumbent (Golosov, 2011; Hale, 2014; Reuter, 2017). United Russia, the party of Vladimir Putin, consolidated its dominant position across successive electoral cycles, and the opposition remained divided along the lines that already separated it (Gel'man, 2013).

Turkey comes next, and it is the most instructive case because it traces the opposition's response across the full range of repression within a single country. Through roughly 2013, Recep Tayyip Erdoğan and the Justice and Development Party (AKP) governed with moderate repression and faced an opposition split three ways, along secularist, Kurdish, and nationalist lines, represented by the Republican People's Party (CHP), the Kurdish political movement organized after 2012 in the Peoples' Democratic Party (HDP), and the Nationalist Movement Party (MHP). These parties drew on separate constituencies and rarely worked together against the incumbent (Esen & Gümüşçü, 2016; Somer, 2016). Repression then intensified, first after the 2013 Gezi protests and sharply after the 2016 coup attempt, through mass dismissals from public service, prosecutions, party closures, and the jailing of Selahattin Demirtaş and other HDP figures (Esen & Gümüşçü, 2019; White & Herzog, 2016). The opposition changed course. The CHP, joined by the Good Party (IYI) and smaller parties, formed the Nation Alliance before the 2018 elections, and the HDP backed the bloc by pulling its candidates in close races. In 2019 this cooperation delivered Istanbul, Ankara, and Izmir to opposition mayors, with Ekrem İmamoğlu ending a quarter century of AKP-aligned control of Istanbul (Esen & Gümüşçü, 2019). The bloc ran Kemal Kılıçdaroğlu as a single challenger to Recep Tayyip Erdoğan in the 2023 presidential election but fell short. By the 2024 local elections the formal coalition had dissolved, yet opposition voters consolidated behind the CHP all the same, and the party came first nationally for the first time since 1977 (Yavuz & Koç, 2024). The incumbent who had gained from a divided opposition under moderate repression saw the opposition draw together once repression passed a certain point.

Venezuela follows as a sharper version of the same pattern. Under Hugo Chávez, repression rose gradually through media restrictions, judicial capture, and the selective prosecution of opposition figures, while elections went on and the opposition stayed legal but divided (Corrales & Penfold, 2011; Levitsky & Loxton, 2013). Under Nicolás Maduro from 2013 onward, repression escalated, with opposition leaders jailed, parties banned, and security forces sent against street protests, all against the backdrop of a severe economic collapse (Corrales, 2015). The divided opposition drew together behind the Democratic Unity Roundtable (MUD), a coalition formed in the late Chávez years that came to span the ideological spectrum and won a two-thirds majority in the 2015 National Assembly elections (Corrales, 2020; Gamboa, 2022). Nicolás Maduro answered by stripping the National Assembly of its powers and convening a parallel Constituent Assembly in 2017, moving the contest outside electoral channels. The escalation illustrates the far end of the range the model describes: repression severe enough that even a united opposition could no longer convert cooperation into a path to power.

We close with Hungary, where the main instruments of incumbent power were not repression in the narrow sense. The case still fits the hybrid regime category, and Viktor Orbán himself called his project an *illiberal democracy*. After 2010, Orbán and the Fidesz–Hungarian Civic Alliance (Fidesz) entrenched their position through constitutional revision, the rewriting of electoral law, the capture of independent media, the steering of public contracts to allied firms, and the building of foreign patronage ties, in what amounted to a steady takeover of state institutions (Bozóki & Hegedűs, 2018; Scheppele, 2022; Szeidl & Szűcs, 2021). These tools differ from direct repression, but they reshape the political environment in the same way the model describes: they raise the cost of mounting an independent challenge and sharpen the sense of a single dominant adversary, which pulls the opposition toward common action. Through the 2010s the opposition stayed divided across a wide ideological range, from the Hungarian Socialist Party and the Democratic Coalition to the formerly far-right Jobbik and the centrist Momentum, each holding its own ground (Krekó & Enyedi, 2018). By the 2022 election, six of these parties ran together under the United for Hungary banner behind a single candidate for prime minister, Péter Márki-Zay, yet Fidesz held its two-thirds majority. The pull toward consolidation went further before the 2026 election, as the fragmented opposition gave way to a single dominant challenger when Péter Magyar's Respect and Freedom Party (Tisza) absorbed much of the anti-Fidesz vote. Tisza won 141 of the 199 seats, a two-thirds majority, ending sixteen years of Fidesz government (Kassam & Garamvolgyi, 2026; OSCE Office for Democratic Institutions and Human Rights, 2026).

The cases share the same broad pattern despite their differences. The opposition stays divided where the political environment leaves room for separate factions to go their own way, and moves toward cooperation as that room narrows. The incumbent's prospects for staying in power move with this response. Fragmented oppositions coincided with secure incumbents in Russia and in the early periods of the other three cases, while consolidated oppositions produced the sharpest setbacks the incumbents faced: the loss of Istanbul, Ankara, and Izmir in 2019 and the nationwide first place of the CHP in 2024, the two-thirds opposition majority in Venezuela's 2015 National Assembly elections, and the end of sixteen years of Fidesz government in 2026. These outcomes reflect many forces beyond opposition structure, and in Hungary the instruments of incumbent control were not repression in the narrow sense. On their own the cases cannot establish a general regularity, and opposition behavior in hybrid regimes does not reduce to a single account. They serve here to bring out the patterns of behavior that motivate the model developed in the next sections.

3 Opposition behavior: Framework

This section presents the theoretical framework. We consider a passive incumbent I and a countably infinite set of potential opposition factions $f \in F$ competing in a political contest. We model this as a dynamic game that unfolds over periods $t = 0, 1, 2, \ldots$, and lasts as long as the incumbent stays in power.

3.1 Basics

Before the game starts, a level of repression $R > 0$ is realized and becomes common knowledge. This level stays fixed throughout the game. Repression captures the state's machinery for raising the price of political activity: surveillance, harassment, arrest, intimidation, and violence directed at those who challenge the incumbent. In the model, R is the per-period cost a faction pays whenever it is politically active. Once R is realized, the game begins, and it proceeds in two layers: a one-shot participation decision and a repeated interaction decision.

In period $t = 0$, all factions simultaneously choose participation decisions $x_f \in \{i, a\}$: each faction either becomes *active*, $x_f = a$, and takes part in the contest or remains *inactive*, $x_f = i$, and stays out of it. These one-shot choices fix the set of active factions $F_a = \{f \in F : x_f = a\}$ and their number $n = \sum_{f \in F} \mathbb{1}_{\{x_f = a\}}$ for the rest of the game.

Active factions engage publicly in opposition politics: they organize supporters, campaign, and present themselves to voters as alternatives to the incumbent.[8] This visibility is what exposes them to repression. Inactive factions stay out of the contest and, by staying invisible, out of the incumbent's reach. Activity therefore carries both an opportunity and a cost. An active faction has the chance of winning the contest and succeeding the incumbent. In return, it pays the repression cost R in every period $t \geq 1$.[9]

In every period $t \geq 1$, all active factions $f \in F_a$ simultaneously choose interaction decisions $y_f^t \in \{c, d\}$: each active faction either chooses to *cooperate*, $y_f^t = c$, by joining the other active factions in a joint challenge to the incumbent, or to *defect*, $y_f^t = d$, by breaking away from that challenge and pursuing an independent path to office. We let $m^t = \sum_{f \in F_a} \mathbb{1}_{\{y_f^t = c\}}$ denote the number of cooperating factions in period t.

Cooperation means adopting a united oppositional stance: a cooperating faction refrains from undermining the other factions, aligns around common positions, and directs its criticism at the incumbent rather than at them. Defection means acting independently: a defecting faction emphasizes its own profile, distances itself from those common positions, and seeks individual advantage even when doing so weakens the opposition.

The choice between cooperation and defection involves a trade-off between collective and individual returns. Cooperation raises the probability that the

[8] Opposition to the incumbent takes many forms beyond organized factions. Independent media, non-governmental organizations, labor unions, religious associations, and youth movements all contest the incumbent's rule and frequently draw repression of their own. We concentrate on opposition politics with factions that can succeed the incumbent, since it is this prospect that shapes their strategic interaction, both whether to become active and whether to cooperate with one another. Press and civil society shape the environment in which this contest unfolds, and we read these actors as part of the environment in which factions decide whether to enter rather than as contestants for the incumbent's office.

[9] Note that the participation decision is one-shot: a faction that becomes active remains active for the rest of the game. We view this as the natural case, since a faction that has revealed itself to the incumbent cannot escape repression by simply stepping back, which aligns with empirical observations of opposition factions in hybrid regimes rarely withdrawing once they have become active (Brownlee, 2007; Bunce & Wolchik, 2011; LeBas, 2011; Levitsky & Way, 2010). Given irreversibility of participation, the assumption that entry happens only once at the beginning of the game is without loss for our results. In our setting, a faction that prefers to be active in some period prefers to be active in every earlier period as well. Hence, allowing a fresh participation decision in each period does not change the equilibrium number of active factions. A distinct alternative is to let entry itself be reversible, so that a faction chooses to be active for each individual period and pays the repression cost only in the periods in which it is active. Under relatively mild restrictions, our results carry over to this setting as well, which we discuss in more detail in Section 6.

opposition unseats the incumbent, but it requires each faction to subordinate its programmatic distinctiveness and public visibility to a shared agenda. Defection lets a faction cultivate a recognizable identity and capture a distinct constituency, yet it fragments the opposition and weakens its electoral leverage. This conflict between what serves the opposition as a whole and what serves each faction on its own is well documented in hybrid regimes: unified oppositions are markedly more successful at unseating incumbents, yet unity often breaks down because individual factions gain from standing apart.[10]

At the end of each period $t \geq 1$, a probabilistic contest selects a winner from among the incumbent and the active factions. The incumbent stays in power with probability P_I^t, and each active faction f wins the contest with probability P_f^t. These probabilities depend on the number of active factions and on their interaction decisions. We specify them in Section 3.2.

If the incumbent loses the contest, the game ends and payoffs are realized. The faction that defeats the incumbent takes office and obtains Π_1, the value of controlling the state. Office goes to one faction alone even when the challenge is joint, so cooperation raises the factions' chances against the incumbent without dividing the prize among them. Every other faction, including those that remained inactive, obtains Π_0, the value of the incumbent's fall to the rest of the opposition: repression ends and the political arena reopens for every faction, not only for the winner. We assume $\Pi_0 \in (0, \Pi_1)$: every faction gains from the incumbent losing power, $\Pi_0 > 0$, and every faction strictly prefers to be the one that takes office, $\Pi_1 > \Pi_0$. The difference $\Delta\Pi = \Pi_1 - \Pi_0$ is therefore the office premium, the part of the prize that only the winner captures and that the factions contest among themselves. If the incumbent wins the contest, no payoffs are realized in that period beyond the repression cost, and the game proceeds to period $t + 1$.

3.2 Probabilities of winning

We now specify the structure of the political contest. The contest follows a Tullock-like structure (Tullock, 1980). Unlike in the standard Tullock contest, however, effort is not chosen endogenously. Instead, the strategic content of the

[10]The strategic problem facing active factions has the structure of a classic collective action problem (Olson, 1965). Opposition coalitions substantially improve the chances of unseating incumbents in competitive authoritarian regimes (Bunce & Wolchik, 2011; Howard & Roessler, 2006; Wahman, 2013), yet such coalitions are difficult to build and to sustain (Arriola et al., 2021; Buckles, 2019; Gamboa, 2022; Hassan, 2024).

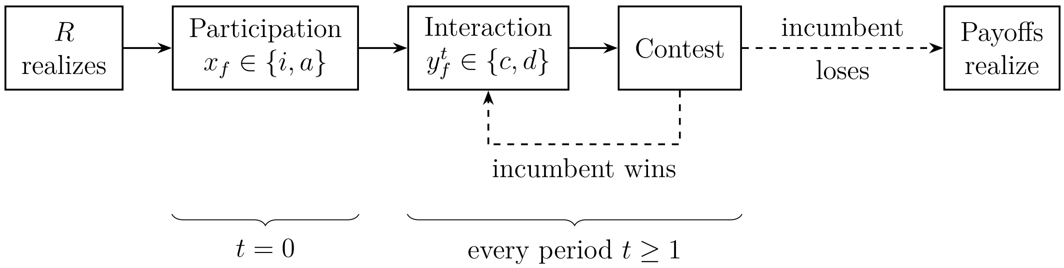


Figure 1: Timing of the model.

model lies entirely in the participation and interaction decisions of opposition factions. Each active faction enters the contest with weight one, while the incumbent enters with a structural advantage $\zeta(n, m^t)$. The active opposition therefore has collective weight n. The probability that the incumbent stays in power in period t is

$$P_I^t = p(n, m^t) = \frac{\zeta(n, m^t)}{\zeta(n, m^t) + n}. \tag{1}$$

This probability falls in n and rises in ζ: a larger active opposition lowers the incumbent's chances, while a greater structural advantage raises them.

The incumbent's structural advantage depends on how the active factions interact. When every active faction cooperates, the opposition confronts voters as a single alternative: public attention concentrates on the incumbent's record rather than on divisions within the opposition, and the incumbent's structural advantage falls.[11] We assume this effect requires full cooperation. A single defecting faction shifts attention away from the incumbent's record and onto the split itself, and the advantage returns to its higher level.[12] Formally, the incumbent's structural advantage is defined as follows:

$$\zeta(n, m^t) = \begin{cases} \zeta_d & \text{if } n > m^t, \\ \\ \zeta_c & \text{if } n = m^t, \end{cases} \tag{2}$$

[11] Voters in hybrid regimes rally around a focal, unified alternative: opposition breakthroughs against incumbents are regularly preceded by the formation of a single coalition or joint candidacy (Bunce & Wolchik, 2011; Howard & Roessler, 2006; Wahman, 2013).

[12] We assume voters respond to visible disunity itself rather than to the motives behind it: a public split is read as opposition incoherence regardless of which faction caused it. Voters rely on coarse public signals such as labels and joint stances when assessing alternatives (Snyder & Ting, 2002), and incumbents in hybrid regimes deliberately cultivate visible opposition splits precisely because such splits weaken the opposition's standing with voters (Arriola et al., 2021; Hassan, 2024).

with $\zeta_c < \zeta_d$.[13]

We now turn to the win probabilities of individual factions. The interaction decisions determine which factions are visible to voters as alternatives to the incumbent. A cooperating faction folds its own profile into the united oppositional stance: it stands before voters as part of the opposition, not as a distinct alternative in its own right. A defecting faction sharpens its own profile at the cooperators' expense: by setting itself apart from the common positions, it presents itself to voters as an alternative not only to the incumbent but also to the rest of the opposition. Conditional on the incumbent losing the contest, office therefore goes to a faction that stands out, since voters can install only a faction they perceive as a distinct alternative to the incumbent. The residual probability of an incumbent defeat is allocated across factions according to

$$\hat{p}(x_f, n, y_f^t, m^t) = \begin{cases} 0 & \text{if } x_f = i, \text{ or if } n > m^t \text{ and } y_f^t = c, \\ \dfrac{1}{n - m^t} & \text{if } n > m^t \text{ and } y_f^t = d, \\ \dfrac{1}{n} & \text{if } n = m^t. \end{cases} \tag{3}$$

An inactive faction never takes office. When the active opposition is unified, $n = m^t$, no faction stands out above the rest and all active factions share the residual probability equally. When some factions defect, $n > m^t$, the defectors alone are visible as alternatives: they divide the residual probability equally among themselves, and the cooperating factions obtain nothing, since a faction that has folded into the united stance does not stand out as a distinct alternative for voters to turn to.

A faction's overall win probability in period t is then the product of its individual share and the probability of an incumbent defeat:

$$P_f^t = \hat{p}(x_f, n, y_f^t, m^t)\,(1 - p(n, m^t)). \tag{4}$$

The factor $1 - p(n, m^t)$ is the probability that the incumbent loses the contest in period t, and $\hat{p}$ distributes this probability across the factions according to their decisions.

This structure embeds both the collective benefit of cooperation and its

[13] This all-or-nothing form can be understood as the limit case of a structural advantage that is decreasing and concave in the number of cooperating factions m^t, as concavity tends to infinity. We discuss this assumption in more detail in Section 6.

individual fragility. A unified opposition lowers the incumbent's structural advantage to ζ_c and gives every cooperating faction a positive chance of taking office. A single defection restores the higher advantage ζ_d and concentrates the entire residual probability among the defectors, leaving the cooperating factions with no chance of taking office.

3.3 Intertemporal payoffs

We now turn to the intertemporal payoff of each faction. Let U_f^t denote the expected payoff of faction f from period t onward. For every period $t \geq 1$ it satisfies the recursion

$$U_f^t = -R\,\mathbb{1}_{\{x_f=a\}} + P_f^t\,\Pi_1 + \sum_{\tilde{f}\neq f} P_{\tilde{f}}^t\,\Pi_0 + P_I^t\,\beta\,U_f^{t+1}. \tag{5}$$

The expression has four terms: the repression cost and one term for each possible outcome of the contest in period t. The first term is the repression cost. Faction f pays R in period t if it is active, $x_f = a$, and pays nothing otherwise. The second and third terms are the contest payoffs in period t. With probability P_f^t, faction f defeats the incumbent and obtains Π_1. With probability $\sum_{\tilde{f}\neq f} P_{\tilde{f}}^t$, some other faction defeats the incumbent and faction f obtains Π_0. The fourth term is the continuation payoff. With probability P_I^t, the incumbent stays in power, the game proceeds to period $t+1$, and faction f's future payoff U_f^{t+1} is discounted by the common discount factor $\beta \in (0,1)$.

3.4 Histories and strategies

We now formalize the histories and strategies of the game. Let $h^0 \in H^0 = \{\emptyset\}$ denote the initial history before any decisions are made. A history at the start of period $t \geq 1$ records the participation decisions of all factions and, for each period s with $1 \leq s < t$, the interaction decisions of the active factions. Let H^t denote the set of all such histories and let $H = \bigcup_{t\geq 0} H^t$. We restrict attention to pure strategies. A strategy of faction f is a tuple $\sigma_f = (x_f, y_f)$, where $x_f \in \{i, a\}$ is the participation decision and, if $x_f = a$,

$$y_f : \bigcup_{t\geq 1} H^t \to \{c, d\} \tag{6}$$

maps any history $h^t \in H^t$ to an interaction decision in period t. A strategy profile is the collection $\sigma = (\sigma_f)_{f\in F}$. Given a strategy profile σ and a history

$h \in H$, the pair (σ, h) defines all relevant objects at any period $\hat{t} \geq 0$: the induced histories $h^{\hat{t}}(\sigma, h)$, the active set $F_a(\sigma, h)$, the numbers of active and cooperating factions $n(\sigma, h)$ and $m^{\hat{t}}(\sigma, h)$, and the continuation value $U_f^{\hat{t}}(\sigma, h)$ for any $f \in F_a(\sigma, h)$. These objects are determined by h alone when they concern decisions before the current node, and recursively by σ and h along the path of play when they refer to future decision nodes.

4 Opposition behavior: Formal analysis

We now analyze equilibrium behavior in the game presented in Section 3. Infinitely repeated games typically admit a large set of equilibria. We restrict attention along two dimensions. First, we focus on pure-strategy subgame perfect Nash equilibria. Second, we restrict attention to equilibria in which strategies condition only on participation decisions: a faction's interaction decision in any period may depend on which factions are active, but not on how active factions have behaved in previous periods.[14] This restriction reflects a basic asymmetry in the observability of behavior in the studied environment: while participation behavior is public and therefore necessarily observable, sustaining equilibria that condition on past interaction behavior requires a degree of monitoring and record-keeping that is often difficult to achieve in hybrid regimes.[15] The following definition states our equilibrium notion formally.

[14]Note that this restriction rules out history-dependent punishment and reward schemes based on past interaction behavior, including grim trigger strategies.

[15]Verifying past interaction behavior is difficult where incumbents manipulate the information environment (Edmond, 2013; Guriev & Treisman, 2020; King et al., 2013; Roberts, 2018) and where campaigning is multidimensional and hard to monitor comprehensively (Schedler, 2013). Moreover, the organizational memory required to sustain strategies based on long histories of past play is often weakened by leadership turnover, repression, and organizational restructuring, limiting the preservation and transmission of information about past interactions (Arriola et al., 2021; Brownlee, 2007; LeBas, 2011; Reuter, 2017).

Definition 1. ***Opposition equilibrium***
A strategy profile σ is an opposition equilibrium if and only if for all $f \in F$, $t \geq 0$, and $h \in H^t$:

1. *for all $\hat{\sigma}$ with $\hat{\sigma}_{\hat{f}} = \sigma_{\hat{f}}$ for all $\hat{f} \neq f$,*

$$U_f^t(\hat{\sigma}, h) \leq U_f^t(\sigma, h),$$

2. *if $t \geq 1$ and $f \in F_a(h)$, then for all $\hat{h} \in H$ that coincide with h on $x_{\hat{f}}$ for all $\hat{f} \in F$,*

$$\sigma_f(\hat{h}) = \sigma_f(h).$$

Our analysis identifies two such opposition equilibria. In the first, all active factions defect and the opposition operates as a fragmented set of independent contenders, each setting itself apart from the rest and drawing its own share of the residual probability of an incumbent defeat. In the second, all active factions cooperate and the opposition confronts voters as a unified front, no faction standing out and all sharing that probability equally.

4.1 Defection equilibrium

We first characterize the opposition equilibrium in which every active faction defects in every period. The analysis proceeds in two steps: the participation decision first, then the interaction decision.

4.1.1 Participation decision

We first determine how many factions become active when every active faction defects in every period. Let $V_i^D(n)$ denote the expected present value of a faction that remains inactive when n active factions all defect. The subscript i marks inactivity, and the superscript D marks universal defection.

A faction that remains inactive pays no repression cost and never takes office. With probability $1 - p(n, 0)$, the incumbent loses the contest and the faction receives Π_0. With probability $p(n, 0)$, the incumbent stays in power and the game continues, with future payoffs discounted by β. The value of inactivity therefore solves

$$V_i^D(n) = (1 - p(n,0))\,\Pi_0 + p(n,0)\,\beta\,V_i^D(n)$$
$$= \frac{n\,\Pi_0}{\zeta_d(1-\beta)+n}. \tag{7}$$

The value $V_i^D(n)$ rises in Π_0 and in β, since a larger payoff from regime change and a higher weight on future payoffs both raise the continuation value of waiting on the sidelines. It falls in ζ_d, since a stronger incumbent is less likely to ever lose power. Finally, $V_i^D(n)$ rises in n at a decreasing rate (see Lemma A.1 in Appendix A.1). Adding active factions raises the probability of an incumbent defeat, but the marginal effect is largest at small n and shrinks as the field grows.

Next, consider the value of being active. Let $V_d^D(n)$ denote the expected present value of a faction that is active and defects in every period when $n-1$ other active factions defect as well.[16] The subscript d marks activity with defection.

An active faction pays the repression cost R. With probability $1-p(n,0)$, the incumbent loses the contest. The faction then receives Π_0, and with probability $1/n$ it is the faction that takes office and earns in addition the office premium $\Delta\Pi$. With probability $p(n,0)$, the incumbent stays in power and the game continues, with future payoffs discounted by β. The value of activity solves

$$V_d^D(n) = -R + (1-p(n,0))\Big(\Pi_0 + \tfrac{\Delta\Pi}{n}\Big) + p(n,0)\,\beta\,V_d^D(n)$$
$$= \frac{\Delta\Pi + n\,\Pi_0 - R\,(\zeta_d + n)}{\zeta_d(1-\beta)+n}. \tag{8}$$

$V_d^D(n)$ rises in Π_1 and weakly rises in Π_0, with the effect of Π_0 strict whenever $n > 1$. Both parameters raise the payoff an active faction expects when the incumbent loses the contest. When $V_d^D(n) > 0$, the value also rises in β, since future payoffs receive greater weight. It falls in ζ_d, which lowers the probability of an incumbent defeat, and in R, which lowers the payoff of activity directly (see Lemma A.2 in Appendix A.1).

The effect of n on $V_d^D(n)$ is ambiguous. A larger field of active factions raises the probability of an incumbent defeat, which benefits a defector through

[16]The value $V_d^D(n)$ is defined for $n \geq 1$, since it presupposes at least one active faction.

Π_0, but it also lowers each defector's chance of taking office, which works against the defector through the office premium. When Π_1 is small relative to Π_0, ζ_d, and R, the second effect is small and the value rises in n. When Π_1 is large relative to these terms, the second effect dominates and the value falls in n (see Lemma A.3 in Appendix A.1).

We can now determine how many factions become active. A currently inactive faction obtains value $V_i^D(n)$ when n other factions are active and defect. By becoming active, it raises the number of active factions to $n+1$ and obtains $V_d^D(n+1)$. Hence, an inactive faction weakly prefers to become active when n factions are already active if and only if

$$\Delta V^D(n) := V_d^D(n+1) - V_i^D(n) \geq 0. \tag{9}$$

We define the threshold at which an inactive faction is indifferent between becoming active and remaining inactive as

$$\tilde{n}^D \geq 0 \text{ such that } \Delta V^D(\tilde{n}^D) = 0, \tag{10}$$

whenever such a value exists. Lemma A.4 in Appendix A.1 establishes that this threshold exists if and only if repression is not too large, $R \leq \Pi_1/(\zeta_d + 1)$, is unique whenever it exists, and that $\Delta V^D(n)$ crosses zero from above at $\tilde{n}^D$. As a consequence, an inactive faction wishes to become active if and only if $n \leq \tilde{n}^D$.

The intuition is as follows. When few factions are active, $n \leq \tilde{n}^D$, the gain from becoming active is large: the probability of winning the contest is high because there are few active factions, and the marginal effect of one more active faction on unseating the incumbent is also large. Both gains shrink as more factions become active, and at $\tilde{n}^D$ these gains are completely offset by the repression costs that activity entails.

Since inactive factions wish to become active as long as $n \leq \tilde{n}^D$, the equilibrium number of active factions under universal defection is the smallest integer strictly larger than $\tilde{n}^D$:

$$n^D := \min\{x \in \mathbb{N} : x > \tilde{n}^D\}, \quad \text{whenever } \tilde{n}^D \geq 0 \text{ exists.} \tag{11}$$

The following lemma summarizes these insights formally.

Lemma 1. ***Defection equilibrium: Participation***

Suppose $\tilde{n}^D$ exists. Let $D_x := \{\sigma : m^t(\sigma, \emptyset) = 0$ for all $t \geq 1\}$. For any $f \in F$, $\sigma \in D_x$ such that $n(\sigma, \emptyset) = n^D$, and $\hat{\sigma} \in D_x$ with $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ for all $\tilde{f} \neq f$,

$$U_f^0(\hat{\sigma}, \emptyset) \leq U_f^0(\sigma, \emptyset).$$

Proof: See Appendix A.1.

The set D_x collects all strategy profiles in which every active faction defects in every period along the path of play, so that $m^t(\sigma, \emptyset) = 0$ for all $t \geq 1$. The lemma takes such a profile with exactly n^D active factions and states that no faction can gain by changing its participation decision, holding the defection behavior fixed. The intuition follows from the threshold logic of Section 4.1.1.

We now turn to how the model's parameters affect n^D. The parameter of primary interest is repression R.

Lemma 2. ***Repression and*** n^D

n^D satisfies the following properties whenever it exists:

2.1 $\lim_{R \to 0^+} n^D = \infty$.

2.2 $\lim_{R \to \Pi_1/(\zeta_d+1)^-} n^D = 1$.

2.3 n^D *is decreasing in* R.

Proof: See Appendix A.1.

The underlying intuition behind Lemma 2 is that higher repression discourages participation by lowering the payoff from activity. Consequently, an increase in repression decreases the number of factions n^D. As repression approaches 0 from above, participating becomes essentially costless in the model and the number n^D goes to infinity. As repression R tends towards $\Pi_1/(\zeta_d + 1)$, only a single faction wants to become active under universal defection.

The number of active factions n^D also responds to the other parameters. It rises in Π_1 and falls in Π_0 and β, and for sufficiently large ζ_d it falls in ζ_d (see Lemma A.7 in Appendix A.1).

4.1.2 Interaction decision

Next, we characterize the conditions under which universal defection is sustainable in equilibrium. A single faction defecting destroys the opposition's standing as a unified alternative, while a cooperator facing defectors forfeits any chance of taking office. A faction therefore defects whenever it expects even one other active faction to do so, making universal defection sustainable whenever at least two factions are active.

A sole active faction, by contrast, optimally cooperates: with no rival to coordinate against, cooperation simply means facing voters under the lower structural advantage ζ_c, so choosing it is mechanical rather than strategic (see Lemma A.9 in Appendix A.1). The following lemma captures these insights formally.

Lemma 3. ***Defection equilibrium: Interaction***
Consider any σ such that for all $t \geq 1$ and $h \in H^t$,

1. $n(\sigma, h) \geq 2 \Longrightarrow m^t(\sigma, h) = 0$ *and*
2. $n(\sigma, h) = 1 \Longrightarrow m^t(\sigma, h) = 1.$

For any $f \in F$, $t \geq 1$, $h \in H^t$, and all $\hat{\sigma} \neq \sigma$ with

1. $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ *for all* $\tilde{f} \neq f$ *and*
2. $\hat{\sigma}_f(\tilde{h}) = \sigma_f(\tilde{h})$ *for all* $\tilde{h} \neq h$

it holds that

$$U_f^t(\hat{\sigma}, h) \leq U_f^t(\sigma, h).$$

Proof: See Appendix A.1.

The lemma states that when at least two factions are active and all defect, or when a sole active faction cooperates, no active faction has an incentive to unilaterally deviate from that interaction behavior in any period $t \geq 1$.

4.1.3 Equilibrium description

The previous results combine to characterize the defection equilibrium and the conditions under which it exists.

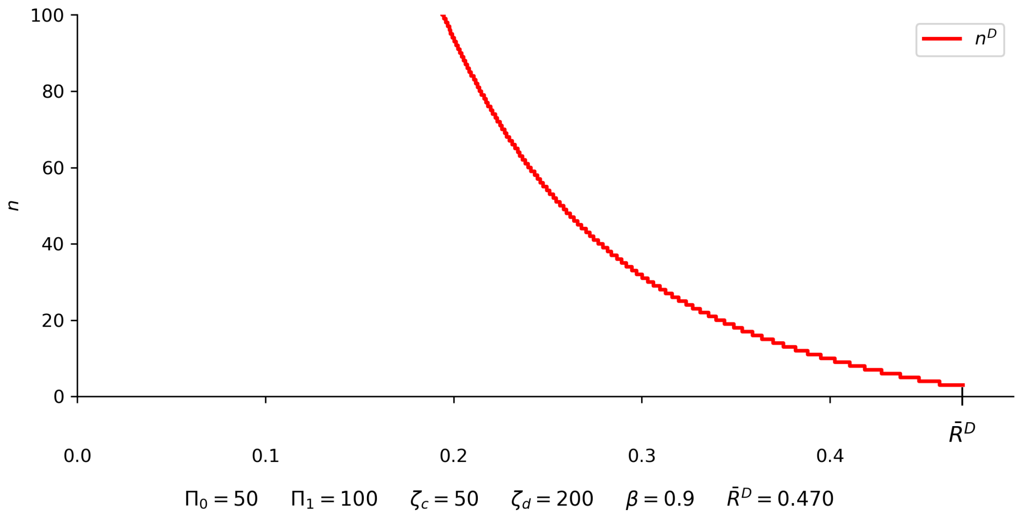


Figure 2: Defection equilibrium.

Proposition 1. ***Defection equilibrium***
Let $D := \{\sigma : n(\sigma, \emptyset) = n^D$ *and* $m^t(\sigma, \emptyset) = 0$ *for all* $t \geq 1\}$. *There exists* $\bar{R}^D > 0$ *such that*

$$R \in (0, \bar{R}^D] \iff D \text{ contains an opposition equilibrium.}$$

Proof: See Appendix A.1.

The proposition states that a defection equilibrium with n^D active factions exists if and only if repression is not too large, that is, $R \leq \bar{R}^D$ for some $\bar{R}^D$. The result follows largely from Lemmas 1 and 3. The intuition is as follows. When repression is not too large, at least two factions want to become active under universal defection, $n^D \geq 2$. Lemma 1 then pins down the equilibrium number of active factions at n^D, and Lemma 3 establishes that, with at least two active factions, defection is a best response for each of them. Figure 2 presents an illustration of the defection equilibrium, with the number of active factions given by the threshold n^D. The equilibrium exists for $R \leq \bar{R}^D$, and within that range n^D falls in R.

4.2 Cooperation equilibrium

We now turn to the equilibrium in which every active faction cooperates in every period, presenting a unified front against the incumbent. The structure of the analysis mirrors the defection case, with the participation decision examined first and the interaction decision second.

4.2.1 Participation decision

The participation analysis follows the same logic as in the defection case, with the incumbent's win probability now given by $p(n, n)$ rather than $p(n, 0)$, which amounts to the incumbent's structural advantage being ζ_c rather than ζ_d. The mathematical structure and underlying intuition is otherwise identical, for which reason we keep the exposition compact. Let $V_i^C(n)$ denote the expected present value of a faction that remains inactive in every period when $n \geq 0$ active factions all cooperate. The superscript C marks universal cooperation, and the subscript i marks inactivity. The present value solves

$$\begin{aligned} V_i^C(n) &= (1 - p(n,n))\, \Pi_0 + p(n,n)\, \beta\, V_i^C(n) \\ &= \frac{n\, \Pi_0}{\zeta_c(1-\beta) + n}. \end{aligned} \tag{12}$$

The value $V_i^C(n)$ rises in Π_0 and in β, and it falls in ζ_c. Finally, $V_i^C(n)$ rises in n at a decreasing rate (see Lemma A.11 in Appendix A.2). Let $V_c^C(n)$ denote the expected present value of a faction that is active and cooperates in every period when $n-1$ other factions are active and cooperate as well. The superscript C marks universal cooperation, and the subscript c marks activity with cooperation. The value solves

$$\begin{aligned} V_c^C(n) &= -R + (1 - p(n,n))\left(\tfrac{\Delta\Pi}{n} + \Pi_0\right) + p(n,n)\, \beta\, V_c^C(n) \\ &= \frac{\Delta\Pi + n\, \Pi_0 - R\,(\zeta_c + n)}{\zeta_c(1-\beta) + n}. \end{aligned} \tag{13}$$

Analogously to $V_d^D(n)$, the value $V_c^C(n)$ rises in Π_1 and weakly rises in Π_0, with the effect of Π_0 strict whenever $n > 1$. It falls in R and in ζ_c. When $V_c^C(n) > 0$, it also rises in β (see Lemma A.12 in Appendix A.2). As before, the effect of n on $V_c^C(n)$ is ambiguous. A larger field of cooperating factions raises the probability of an incumbent defeat, but it also lowers each faction's chance of taking office. When Π_1 is small relative to Π_0, ζ_c, and R, the second effect is limited and the value rises in n. When Π_1 is large relative to these terms, the second effect dominates and the value falls in n (see Lemma A.13 in Appendix A.2). When n factions are active and all factions always cooperate, a currently inactive faction weakly prefers to become active if and only if

$$\Delta V_a^C(n) := V_c^C(n+1) - V_i^C(n) \geq 0. \tag{14}$$

We define the threshold at which an inactive faction is indifferent between becoming active and remaining inactive as

$$\tilde{n}_a^C \geq 0 \text{ such that } \Delta V_a^C(\tilde{n}_a^C) = 0, \tag{15}$$

whenever such a value exists. By analogous reasoning as for $\tilde{n}^D$ in Subsection 4.1.1, the threshold $\tilde{n}_a^C$ exists if and only if $R \leq \Pi_1/(\zeta_c + 1)$, is unique whenever it exists, and $\Delta V_a^C(n)$ crosses zero from above at $\tilde{n}_a^C$ (see Lemma A.14 in Appendix A.2). Hence, inactive factions weakly prefer to become active if and only if the number of active factions satisfies $n \leq \tilde{n}_a^C$. The equilibrium number of active factions under universal cooperation is the smallest integer larger than $\tilde{n}_a^C$:

$$n_a^C := \min\{x \in \mathbb{N} : x > \tilde{n}_a^C\}, \quad \text{whenever } \tilde{n}_a^C \geq 0 \text{ exists.} \tag{16}$$

The following lemma states these properties formally. In particular, it states that if all factions always cooperate, then no faction has an incentive to change its participation decision if and only if n_a^C factions become active.

Lemma 4. ***Cooperation equilibrium: Participation***
Suppose $\tilde{n}_a^C$ exists. Let $C_x := \{\sigma : m^t(\sigma, \emptyset) = n(\sigma, \emptyset) \text{ for all } t \geq 1\}$. For any $f \in F$, $\sigma \in C_x$ such that $n(\sigma, \emptyset) = n_a^C$, and $\hat{\sigma} \in C_x$ with $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ for all $\tilde{f} \neq f$,

$$U_f^0(\hat{\sigma}, \emptyset) \leq U_f^0(\sigma, \emptyset).$$

Proof: See Appendix A.2.

The main comparative static of n_a^C is again the effect of repression R.

Lemma 5. ***Repression and*** n_a^C

n_a^C satisfies the following properties whenever it exists:

5.1 $\lim_{R\to 0^+} n_a^C = \infty$.

5.2 $\lim_{R\to \Pi_1/(\zeta_c+1)^-} n_a^C = 1$.

5.3 n_a^C is decreasing in R.

Proof: See Appendix A.2.

The logic mirrors the defection case: higher repression lowers the payoff from activity, which reduces the number of factions that become active. For low levels of repression, as R approaches zero, the number of active factions goes to infinity, while as repression moves toward the upper bound for which n_a^C is defined, that is, as R approaches $\frac{\Pi_1}{\zeta_c+1}$, only a single faction wishes to be active. The threshold n_a^C also rises in Π_1 and falls in Π_0 and β, and for sufficiently large ζ_c it falls in ζ_c (see Lemma A.17 in Appendix A.2).

4.2.2 Interaction decision

We now turn to the interaction decision of active factions. For this purpose, we suppose n factions are active in any given period and analyze when no faction has an incentive to deviate from cooperation. The value of being active and cooperating in every period t when $n-1$ other factions behave so as well was derived above as

$$V_c^C(n) = \frac{\Delta\Pi + n\,\Pi_0 - R\,(\zeta_c + n)}{\zeta_c(1-\beta) + n}. \tag{17}$$

Let $V_d^C(n)$ denote the expected present value of a faction that is active and defects in every period when $n-1$ other active factions cooperate.[17] The defecting faction pays the repression cost R. By defecting, it breaks the unified front, and the probability that the incumbent loses the contest falls to $1-p(n,0)$. When the incumbent loses the contest, the defecting faction wins with certainty, since all other active factions cooperate and forfeit their visibility, so it obtains Π_1. With probability $p(n,0)$, the incumbent stays in power and the game

[17] The value $V_d^C(n)$ is defined for $n \geq 1$, since it presupposes at least one active faction.

continues, with future payoffs discounted by β. The value of defection solves

$$\begin{aligned} V_d^C(n) &= -R + (1 - p(n,0))\,\Pi_1 + p(n,0)\,\beta\,V_d^C(n) \\ &= \frac{n\,\Pi_1 - R\,(\zeta_d + n)}{\zeta_d(1-\beta) + n}. \end{aligned} \tag{18}$$

The value $V_d^C(n)$ rises in Π_1, since a larger payoff from taking office raises the return to defection. When $V_d^C(n) > 0$, it also rises in β, since future payoffs receive greater weight. It falls in ζ_d, which lowers the probability of an incumbent defeat, and in R, which lowers the payoff of defection directly. Finally, $V_d^C(n)$ rises in n at a decreasing rate (see Lemma A.19 in Appendix A.2). A larger field of active factions raises the probability that the incumbent loses the contest, which the sole defector captures fully because all other active factions forfeit their individual visibility by cooperating. To see when full cooperation can be sustained as an equilibrium, suppose n factions are active and all cooperate. A cooperating faction has no incentive to switch to defection if and only if

$$\Delta V_c^C(n) := V_c^C(n) - V_d^C(n) \geq 0. \tag{19}$$

We define the threshold at which an active faction is indifferent between cooperation and defection as

$$\tilde{n}_c^C \geq 0 \text{ such that } \Delta V_c^C(\tilde{n}_c^C) = 0. \tag{20}$$

The threshold always exists and is unique, and $\Delta V_c^C(n)$ crosses zero from above at $\tilde{n}_c^C$. Cooperation is therefore sustainable if and only if $n \leq \tilde{n}_c^C$ (see Lemma A.20 in Appendix A.2). The logic is as follows. When few factions are active, $n \leq \tilde{n}_c^C$, the incumbent faces limited opposition and is likely to remain in power. Cooperation then raises the joint probability of defeating the incumbent, giving every active faction a meaningful chance at office. With few rival opposition factions, the probability of winning conditional on the incumbent losing is relatively high, so competition within the opposition stays mild. Together, a large benefit from cooperating and a small cost of foregoing a unilateral edge make universal cooperation feasible. As the field grows, both forces work against cooperation: the probability of being the sole winner conditional on the incumbent losing falls, increasing the benefit of defection, while a larger field independently raises the probability of unseating the incumbent even without universal cooperation. Once $n > \tilde{n}_c^C$, each faction prefers to defect unilaterally, trading the collective benefit of a unified front for a guaranteed edge over the other opposition factions. Consider, for example, the case where the number of active factions grows without bound: the incumbent loses regardless of cooperation, so every faction strictly prefers to defect in order to get ahead of the other $n-1$ factions. The equilibrium number of active factions under universal cooperation is the largest integer that does not exceed $\tilde{n}_c^C$:

$$n_c^C := \max\{x \in \mathbb{N} : x \leq \tilde{n}_c^C\}. \tag{21}$$

Hence, cooperation is sustainable in an equilibrium if at most n_c^C factions are active. The following lemma captures the above insights formally.

Lemma 6. ***Cooperation equilibrium: Interaction***
Consider any σ such that for all $t \geq 1$ and $h \in H^t$,

1. $n(\sigma, h) \leq n_c^C \Longrightarrow m^t(\sigma, h) = n(\sigma, h)$ *and*
2. $n(\sigma, h) > n_c^C \Longrightarrow m^t(\sigma, h) = 0.$

For any $f \in F$, $t \geq 1$, $h \in H^t$, and all $\hat{\sigma} \neq \sigma$ with

1. $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ *for all* $\tilde{f} \neq f$ *and*
2. $\hat{\sigma}_f(\tilde{h}) = \sigma_f(\tilde{h})$ *for all* $\tilde{h} \neq h$

it holds that

$$U_f^t(\hat{\sigma}, h) \leq U_f^t(\sigma, h).$$

Proof: See Appendix A.2.

The lemma states that factions have no incentive to deviate from their interaction behavior if all active factions always cooperate when at most n_c^C factions are active and all active factions always defect if more than n_c^C factions are active. We continue with how changes in the model parameters affect the cooperation bound n_c^C. The main variable of interest is again repression R.

Lemma 7. ***Repression and*** n_c^C
n_c^C satisfies the following properties:

7.1 $\lim_{R\to 0^+} n_c^C \geq 1$.

7.2 $\lim_{R\to\infty} n_c^C = \infty$.

7.3 n_c^C is increasing in R.

Proof: See Appendix A.2.

The lemma establishes that as repression R approaches 0, the threshold n_c^C is above 1, while as repression R grows without bound, the threshold n_c^C grows without bound as well. For all repression values in between, $R \in (0, \infty)$, the threshold n_c^C increases in repression. The underlying intuition is that an increase in repression increases the benefit of unseating the incumbent in order to avoid this repression in future rounds, which favors cooperation and loosens the cooperation constraint. The threshold n_c^C also responds to the remaining parameters. Within the parameter space relevant for the cooperation equilibrium, it rises in Π_0 and ζ_d and falls in Π_1, ζ_c, and β (see Lemma A.23 in Appendix A.2).[18]

4.2.3 Equilibrium description

The previous two subsections combine to characterize the cooperation equilibrium. The participation analysis established that the largest number of factions willing to be active under universal cooperation is n_a^C. The interaction analysis established that the largest number of factions willing to cooperate when all others do is n_c^C. A cooperation equilibrium with n^C active factions requires both constraints to hold: no active faction prefers to be inactive, $n^C \leq n_a^C$, and no active faction prefers to deviate to defection, $n^C \leq n_c^C$. We focus on the most cooperative such outcome, in the sense that it constitutes the largest number of factions that cooperate each round, given by

$$n^C := \min\{n_a^C, n_c^C\}. \tag{22}$$

[18]The relevant parameter space is the set of parameters for which $V_c^C(\tilde{n}_c^C) = V_d^C(\tilde{n}_c^C) \geq 0$, since otherwise the cooperation threshold $\tilde{n}_c^C$ ceases to be a binding constraint in equilibrium, as will become more apparent in the next section.

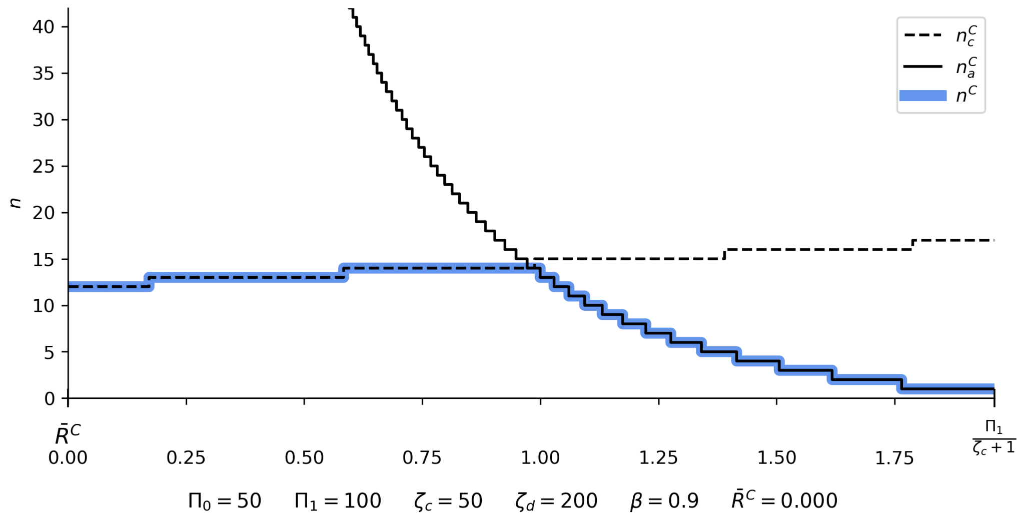


Figure 3: Cooperation equilibrium.

Hence, our cooperation equilibrium is characterized by n^C active factions who cooperate in each round $t \geq 1$. The following proposition characterizes the cooperation equilibrium and the conditions under which it exists.

Proposition 2. ***Cooperation equilibrium***
Let $C := \{\sigma : n(\sigma, \emptyset) = m^t(\sigma, \emptyset) = n^C$ *for all* $t \geq 1\}$. *There exists* $\bar{R}^C < \frac{\Pi_1}{\zeta_c+1}$ *such that*

$$R \in \left[\bar{R}^C, \frac{\Pi_1}{\zeta_c + 1}\right] \iff C \text{ contains an opposition equilibrium.}$$

Proof: See Appendix A.2.

The proposition states that the cooperation equilibrium exists when repression is neither too low nor too high, $R \in [\bar{R}^C, \frac{\Pi_1}{\zeta_c+1}]$. Figure 3 provides a graphical illustration, where the blue line traces n^C over the range of repression values for which the cooperation equilibrium exists. Recall that n_c^C, represented by the dashed line, increases in repression R, while n_a^C, represented by the solid black line, decreases in R, so that the two cross at some repression level. Beginning at the highest repression levels, when $R > \frac{\Pi_1}{\zeta_c+1}$, no faction finds participation worthwhile even under full cooperation. With no active factions, cooperation between factions is vacuous, and the cooperation equilibrium does not exist.[19] As repression falls into the range $R \in (\check{R}, \frac{\Pi_1}{\zeta_c+1}]$, factions begin to

[19]Equivalently, the cooperation equilibrium has zero active factions.

find participation worthwhile, while the number wishing to be active remains small enough that all of them can sustain cooperation, $n_a^C < n_c^C$, and the cooperation equilibrium exists with $n^C = n_a^C$ active factions. As repression falls further, participation becomes sufficiently attractive that more factions would wish to be active than can jointly sustain cooperation, $n_a^C > n_c^C$. The cooperation equilibrium is sustained instead by having n_c^C active factions, with the remaining factions staying inactive to preserve the incentive to cooperate among those that are active.

In the example illustrated in Figure 3, this incentive survives all the way down to $R = 0$, so $\bar{R}^C = 0$ and the cooperation equilibrium exists at every repression level below $\frac{\Pi_1}{\zeta_c+1}$. This need not hold in general. For some specifications of the model, the lower existence bound of repression $\bar{R}^C$ may be strictly positive.[20] When $\bar{R}^C$ is strictly positive, repression falling below it makes the incentive to remain inactive to preserve cooperation too small: when n_c^C factions are active and cooperating, an inactive faction prefers to participate even though doing so breaks cooperation among all active factions. The cooperation equilibrium therefore fails to exist when $R < \bar{R}^C$.

It follows from the above that n^C is potentially non-monotone in repression. For low to intermediate levels, $n^C = n_c^C$ increases with R, since higher repression relaxes the cooperation constraint and allows more factions to sustain cooperation. For intermediate to high levels, $n^C = n_a^C$ decreases with R, as fewer factions find participation worthwhile. The following lemma captures this formally.

Lemma 8. ***Non-monotonicity in the cooperation equilibrium***
There exists $\check{R} \in (\bar{R}^C, \frac{\Pi_1}{\zeta_c+1})$ *such that:*

8.1 For $R \in (\bar{R}^C, \check{R})$*,* n^C *is increasing in* R*.*

8.2 For $R \in (\check{R}, \frac{\Pi_1}{\zeta_c+1})$*,* n^C *is decreasing in* R*.*

Proof: See Appendix A.2.

5 The incumbent's choice of repression

Having analyzed the behavior of opposition factions, we now turn to the incumbent's choice of repression R. For this purpose, we investigate the incumbent's

[20] See Appendix B for some examples.

winning probability as a function of repression, taking the behavior of opposition factions as given by the equilibria characterized in the previous section.[21] We denote the incumbent's winning probability in each period by $P_I(R)$. We suppose that for any repression level R, opposition factions coordinate on the equilibrium that minimizes the incumbent's winning probability among those equilibria that exist.[22] If neither of the two opposition equilibria exists at some R, then the incumbent's probability of winning the contest is one, as no faction becomes active in that case. Formally,

$$P_I(R) := \min_{q \in \mathcal{P}(R) \cup \{1\}} q, \tag{23}$$

where $\mathcal{P}(R) \subseteq \{p(n^D, 0), p(n^C, n^C)\}$ denotes the set of winning probabilities arising from the equilibria that exist at R.[23] Throughout the analysis, we assume repression cannot be chosen arbitrarily but is bounded above by some R^{max}. The bound describes which repression levels are available to the incumbent. Repression is carried out by a coercive apparatus whose reach is fixed by the state's resources and by what its security agents are willing to execute, so repression beyond this reach is simply not available (Svolik, 2012; Tyson, 2018). In hybrid regimes, the bound also follows from the regime type itself: these regimes are defined by formally democratic institutions and a degree of open contestation, and repression levels that restrict this contestation too severely lie outside the setting we study (Levitsky & Way, 2010; Schedler, 2013).[24]

5.1 Formal analysis

In the following, we establish that in some cases the incumbent's optimal repression level is strictly below the maximum feasible level R^{max}. The key mechanism is a possible non-monotonicity of $P_I(R)$: if $P_I(R)$ is decreasing below R^{max}, the global maximum of $P_I(R)$ on $(0, R^{max}]$ may lie strictly below R^{max},

[21]We refrain from modeling the incumbent as an active player to avoid cluttering the notation, but the results would coincide.

[22]This assumption is motivated by two considerations. First, it captures the incumbent's perspective from a max-min standpoint, considering the role of repression in the worst-case equilibrium selection. Second, the equilibrium with the lowest incumbent's winning probability also corresponds to the equilibrium that maximizes the utility of inactive factions, which in turn tends to make it the equilibrium preferred by opposition factions.

[23]That is, $p(n^D, 0) \in \mathcal{P}(R) \Leftrightarrow R \leq \bar{R}^D$ and $p(n^C, n^C) \in \mathcal{P}(R) \Leftrightarrow R \in [\bar{R}^C, \frac{\Pi_1}{\zeta_c + 1}]$.

[24]Instead of an upper bound of feasible repression levels, we could also assume some increasing costs of repression. This would, however, not affect our main argument, as we discuss in more detail in Section 6.

making it suboptimal for the incumbent to repress as much as possible. This non-monotonicity captures the main insight of our analysis and is formalized in the following proposition.

Proposition 3. ***Non-monotonicity of winning probability***
If ζ_c is sufficiently small, ζ_d is sufficiently large, and β is sufficiently large, then the incumbent's winning probability $P_I(R)$ is a non-monotone function of repression R.

Proof: See Appendix A.3.

To illustrate the intuition behind Proposition 3, we employ Figure 4.[25] The top graph depicts the number of active factions in the defection and cooperation equilibria, n^D and n^C, as functions of repression R, each defined over the range of R for which the respective equilibrium exists. This graph merges Figure 2 and Figure 3 into a single picture. The bottom graph maps the number of active factions into the incumbent's winning probability in each equilibrium. The yellow line traces the lower of the two probabilities, and thus indicates the incumbent's winning probability $P_I(R)$ for any given repression level R.

For low levels of repression R, the defection equilibrium features a large number of active factions n^D, since entry is essentially costless: as R approaches zero, all factions have an incentive to become active. The cooperation equilibrium, by contrast, sustains only a small number of active factions n^C, since this number is bounded by how many factions can be simultaneously active while still preserving the incentive to cooperate, that is, $n^C = n_c^C$.[26] The large number of competitors in the defection equilibrium relative to the small number in the cooperation equilibrium implies that the incumbent's winning probability is lower in the defection equilibrium than in the cooperation equilibrium,

[25]Figure 4 presents a specific numerical example of our model. Appendix B provides additional robustness simulations across a range of parameter values, showing that non-monotonicity of $P_I(R)$ arises for a broad set of parameter regions, and not merely for the specific example depicted here.

[26]Note that the cooperation equilibrium may still exist despite a small number of active factions n^C, since for an inactive faction becoming active at n^C entails destroying the cooperation benefit, which in turn drastically increases the incumbent's probability of winning the contest. Hence, inactive factions may prefer to remain inactive even when n^C is small, in order to preserve cooperation. In such cases, a faction would prefer to become active only if other inactive factions did so as well, but not if it were the sole additional entrant beyond the n^C already active factions.

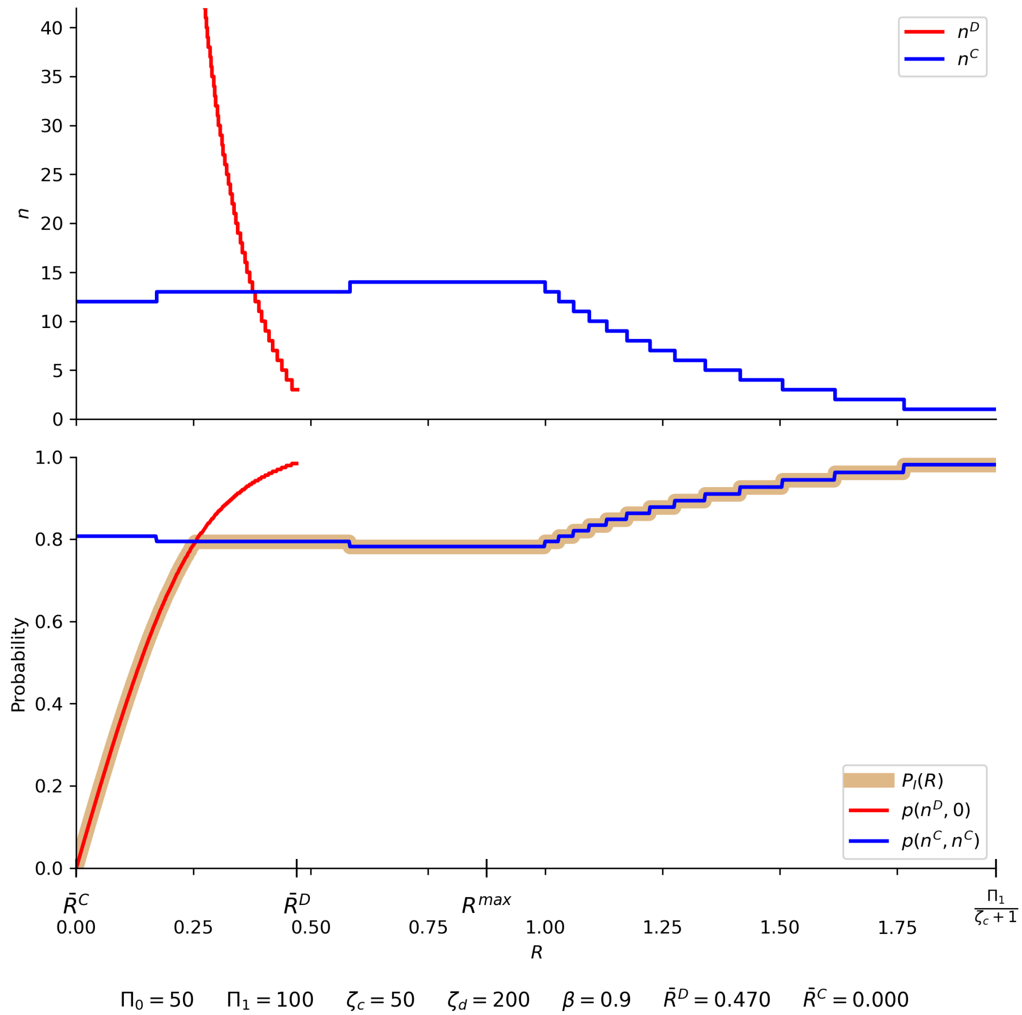


Figure 4: Incumbent's winning probability and repression.

$p(n^D, 0) < p(n^C, n^C)$ (see Lemma A.28 in Appendix A.3).[27]

Given that opposition factions coordinate on the defection equilibrium, an increase in repression R in this region decreases the number of active factions n^D and consequently increases the incumbent's winning probability (see Lemma A.31 in Appendix A.3). In terms of Figure 4, for sufficiently small R, the incumbent's winning probability is given by $P_I(R) = p(n^D, 0)$. Since n^D in the top graph of Figure 4 decreases in R, this maps to an increasing winning probability $P_I(R) = p(n^D, 0)$ in R in the bottom graph.

As repression R increases, the number of active factions in the defection equilibrium n^D decreases, while the number of factions for whom cooperation

[27]For some specifications of the model, there is a region where the cooperation equilibrium does not exist for very small values of R, that is, $\bar{R}^C > 0$.

can be sustained, and thus the number of active factions in the cooperation equilibrium n^C, increases. Eventually, the incumbent faces a lower winning probability in the cooperation equilibrium than in the defection equilibrium, as indicated by the crossing of $p(n^D, 0)$ and $p(n^C, n^C)$ in Figure 4. For repression levels R above this intersection, opposition factions coordinate on the cooperation equilibrium. If repression is not too large, the number of active factions in the cooperation equilibrium remains governed by how many factions can sustain cooperation, $n^C = n_c^C$. Given that opposition factions coordinate on the cooperation equilibrium and the number of active factions is governed by how many can sustain cooperation, an increase in repression R in this region increases the number of active factions n^C and consequently decreases the incumbent's winning probability (see Lemma A.31 in Appendix A.3). Graphically, this corresponds to an increase in R along the rising segment of n^C in Figure 4, which maps to a decreasing winning probability $P_I(R) = p(n^C, n^C)$.

For high levels of repression R, the number of active factions in the cooperation equilibrium is no longer governed by how many factions can sustain cooperation, but instead by how many wish to be active under cooperation, $n^C = n_a^C$. Given that opposition factions coordinate on the cooperation equilibrium and the number of active factions is governed by how many want to be active, an increase in repression R in this region decreases the number of active factions n^C and consequently increases the incumbent's winning probability (see Lemma A.31 in Appendix A.3). Graphically, this corresponds to the values of R for which n^C is declining in the top graph of Figure 4, which maps to an increasing winning probability $P_I(R)$ in R in the bottom graph.

Thus, for low and high levels of repression, the number of active factions is determined by how many factions wish to become active. An increase in repression makes being active more costly, thereby deterring entry of opposition factions and increasing the incumbent's winning probability. For intermediate values of repression, however, the number of active factions that the incumbent faces is determined by how many factions can sustain cooperation. Here, an increase in repression introduces a unifying motivation, namely the avoidance of future repression, which increases the number of active factions that the incumbent faces, and thus decreases the incumbent's winning probability. Consequently, the non-monotonicity of $P_I(R)$ in Figure 4 arises.

Note that Proposition 3 restricts non-monotonicity to some parameter specifications. The underlying reason is that while $P_I(R)$ always exhibits increasing segments, it need not exhibit decreasing segments, since the integer nature of the model may prevent an increase in R from increasing n_c^C on the region where $P_I(R)$ is determined by $p(n^C, n^C)$ and $n^C = n_c^C$. The underlying economic tendency, however, is present for any parameter values: repression always increases factions' incentives to cooperate. The parameter specifications that realize the non-monotonicity are those for which n_c^C is sufficiently responsive to R, and the region over which $P_I(R)$ is determined by $p(n^C, n^C)$ and $n^C = n_c^C$ is sufficiently large.[28]

Non-monotonicity of $P_I(R)$ has a direct implication for the incumbent's optimal choice of repression: repression at its maximum feasible level need not maximize the incumbent's winning probability $P_I(R)$, and a lower level of repression may yield a strictly higher winning probability. Consider R^{max} in Figure 4. The repression level at the intersection of $p(n^D, 0)$ and $p(n^C, n^C)$ yields a strictly larger incumbent's winning probability $P_I(R)$ than R^{max} does. More generally, this logic applies whenever R^{max} falls within an intermediate range in which the non-monotonicity is not yet outweighed by the subsequent rise of $P_I(R)$. If R^{max} is instead very small or very large, maximum repression remains optimal. The following lemma formalizes this.

Lemma 9. ***Optimal level of repression***
If $P_I(R)$ is non-monotone, then there exists R^{max} such that

$$R^{max} \notin \arg \max_{R \in (0, R^{max}]} P_I(R).$$

Proof: See Appendix A.3.

6 Discussion

This paper has examined how repression shapes the behavior of opposition factions in hybrid regimes and, through that behavior, the incumbent's prospects of staying in power. Factions respond to repression through two decisions, whether to become politically active and whether to cooperate with the other active factions. The interplay of the two decisions can render the incumbent's winning probability non-monotone in repression, so that maximal repression

[28] For example, if β is large, then an increase in repression yields a greater benefit from unseating the incumbent to avoid future repression, which makes n_c^C more responsive to R. Moreover, a larger ζ_d decreases the number of active factions n^D for any repression level R, which increases the domain over which $P_I(R)$ is determined by the cooperation equilibrium.

can be strictly suboptimal even though repression carries no direct cost in the model. In what follows, we discuss the main modeling choices behind these results and how the results extend beyond them.

The incumbent enters the analysis only through the level of repression, which is fixed throughout the game and thus equivalent to a choice made once at the start, since Section 5 characterizes the winning probability at every level. A richer variant would let the incumbent adjust repression in every period in response to the observed structure of the opposition. Such adjustment raises a commitment problem: participation decisions are forward looking, so factions weigh the entire anticipated path of repression, and an incumbent unable to commit to future levels may wish to lower repression once a cooperative opposition has formed, which factions anticipate. A full analysis requires modeling the incumbent's strategy explicitly and lies beyond the scope of this paper, but the trade-off at the heart of our analysis, deterring participation while unifying the factions that remain, operates at any repression level the incumbent maintains, so we expect the central tension to survive.

The model also treats participation as permanent: becoming active is a one-time decision, and active factions remain exposed to repression in every subsequent period, capturing the idea from the introduction that entering opposition politics makes a faction publicly identifiable in a way that cannot be undone. Two relaxations deserve mention. First, factions might be able to withdraw from political activity and thereby escape repression. Our results continue to hold when attention is restricted to stationary equilibria. The underlying reason is that under stationarity the same factions are active in every period, so each faction effectively makes a single participation decision, as in the baseline model. Since repression is constant over time, a faction faces the same strategic environment in every period. The model therefore provides no obvious incentive to alternate between periods of activity and inactivity, and stationarity imposes little additional restriction. Stationarity can also be motivated by features we abstract from, such as a fixed cost of becoming active for the first time or an electoral benefit from prior participation. Second, factions might be able to enter in any period rather than only at the outset. Per-period entry alters the incentive of inactive factions to stay out in order to preserve cooperation, the force that pins down the lower bound $\bar{R}^C$ of the cooperation equilibrium. The exact existence range of the equilibrium would therefore change, while the two decisions through which repression operates would not.

The analysis likewise abstracts from any direct cost of repressing and instead imposes an upper bound R^{max} on feasible repression. If repression instead carries increasing costs for the incumbent, for example because a larger coercive apparatus absorbs resources and deepens agency problems, the non-monotonicity of $P_I(R)$ provides a reason for low repression levels even without any binding feasibility constraint. On the decreasing segment of $P_I(R)$, additional repression raises costs and lowers the winning probability at once, so no optimizing incumbent locates there, and reaching the high repression levels at which $P_I(R)$ rises again requires paying the costs of passing through this region, which makes a low repression level attractive whenever costs rise sufficiently steeply. The mechanism we identify thus rationalizes moderate repression under both formulations, a hard capacity bound and increasing costs.

The assumption that opposition factions coordinate on the equilibrium that minimizes the incumbent's winning probability is likewise contestable. Abandoning it, however, gives rise to an additional channel through which larger repression may decrease the incumbent's winning probability, a shift from one equilibrium to another. Consider Figure 4 and suppose opposition factions coordinate on the defection equilibrium whenever it exists. At repression level $\bar{R}^D$, a small increase in repression causes the defection equilibrium to vanish, leaving the cooperation equilibrium as the only equilibrium opposition factions can coordinate on, so the incumbent's winning probability falls from $p(n^D, 0)$ to $p(n^C, n^C)$.[29] Under this alternative equilibrium selection rule, a marginal increase in repression at $\bar{R}^D$ thus causes a discrete decrease in the incumbent's winning probability, a further source of non-monotonicity.

Turning from modeling choices to the results themselves, Proposition 3 establishes non-monotonicity under sufficient conditions without conveying how restrictive these conditions are. The possibility that the winning probability fails to decrease anywhere is entirely due to the integer nature of the model. An increase in repression always strengthens the incentive to cooperate, since the continuous threshold $\tilde{n}_c^C$ is strictly increasing in repression for all parameter values, but the tendency becomes visible in equilibrium only when the increase

[29] Lemma A.32 in Appendix A.3 shows that at repression level $\bar{R}^D$, the winning probability in the defection equilibrium is always larger than that in the cooperation equilibrium, provided the advantage under defection ζ_d is sufficiently larger than the advantage under cooperation ζ_c.

pushes the threshold past the next integer, so the model exhibits threshold effects rather than a smooth response. The unifying tendency of repression is thus always present, and the conditions of the proposition ensure only that it materializes in the equilibrium number of cooperating factions. The simulations in Appendix B document that the non-monotonicity arises across a broad range of parameter specifications, so the result does not depend on a narrow set of parameter values.

Finally, the cooperation we model can be read through the lens of joint candidacies. Cooperating factions forfeit their individual visibility, contest the incumbent as a unified alternative, and one of them takes office when the incumbent loses, a stylized description of opposition alliances that field a joint candidate. An explicit extension would endogenize which faction carries the joint banner. Our symmetric assignment, under which each cooperating faction takes office with equal probability, is the natural benchmark, and asymmetric selection protocols would tighten the cooperation constraint for factions unlikely to lead the alliance, while repression would still raise every faction's benefit from unseating the incumbent, so we expect the unifying channel to persist.

7 Conclusion

The repressive capacity available to incumbents in hybrid regimes is rarely used in full. This paper has argued that such restraint can be strategic. Repression does more than deter: the same instrument that raises the cost of political activity also gives the factions that are active a shared reason to unite against the incumbent, and a united opposition is precisely what an incumbent has most to fear.

In a dynamic model of opposition behavior, we find that the incumbent's winning probability is non-monotone in repression whenever unity genuinely threatens the incumbent and factions care enough about the future. Under these conditions, moderate repression thins a fragmented opposition and works as intended, while pushing further unites the factions that remain, so that beyond a point additional repression lowers the incumbent's chances of staying in power. Repressing as much as possible can then be strictly suboptimal, and an incumbent maximizing the probability of staying in power deliberately holds back.

The central lesson is that the strength of an opposition lies not only in its size but also in its structure. A large and fragmented opposition can leave the incumbent more secure than a small and united one, and repression moves the opposition along both dimensions at once. Observed moderation in repression is therefore not necessarily a sign of limited capacity or normative commitment. It can be the deliberate choice of an incumbent avoiding the unified opposition that harsher repression would create.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the authors used *Claude* and *ChatGPT* to improve language and readability, help structure and refine arguments, support brainstorming, and obtain a preliminary impression of relevant literature. After using these tools, the authors reviewed and edited the text as needed and take full responsibility for it.

References

Acemoglu, D., Egorov, G., & Sonin, K. (2008). Coalition formation in non-democracies. *The Review of Economic Studies*, *75*(4), 987–1009.

Acemoglu, D., Gitmez, A. A., & Shadmehr, M. (2026, June). *Automation and repression* [NBER Working Paper No. 35336].

Acemoglu, D., & Robinson, J. A. (2005). *Economic origins of dictatorship and democracy*. Cambridge University Press.

Arriola, L. R., DeVaro, J., & Meng, A. (2021). Democratic subversion: Elite cooptation and opposition fragmentation. *American Political Science Review*, *115*(4), 1358–1372.

Boix, C., & Svolik, M. W. (2013). The foundations of limited authoritarian government: Institutions, commitment, and power-sharing in dictatorships. *The Journal of Politics*, *75*(2), 300–316.

Bove, V., Platteau, J.-P., & Sekeris, P. G. (2017). Political repression in autocratic regimes. *Journal of Comparative Economics*, *45*(2), 410–428.

Bozóki, A., & Hegedűs, D. (2018). An externally constrained hybrid regime: Hungary in the European Union. *Democratization*, *25*(7), 1173–1189.

Brownlee, J. (2007). *Authoritarianism in an age of democratization*. Cambridge University Press.

Buckles, G. T. (2019). Internal opposition dynamics and restraints on authoritarian control. *British Journal of Political Science*, *49*(3), 883–900.

Bueno de Mesquita, B., Smith, A., Siverson, R. M., & Morrow, J. D. (2003). *The logic of political survival*. MIT Press.

Bueno de Mesquita, E., & Shadmehr, M. (2023). Rebel motivations and repression. *American Political Science Review*, *117*(2), 734–750.

Bunce, V. J., & Wolchik, S. L. (2011). *Defeating authoritarian leaders in postcommunist countries*. Cambridge University Press.

Chiang, A. Y. (2021). Violence, non-violence and the conditional effect of repression on subsequent dissident mobilization. *Conflict Management and Peace Science*, *38*(6), 627–653.

Corrales, J. (2015). The authoritarian resurgence: Autocratic legalism in Venezuela. *Journal of Democracy*, *26*(2), 37–51.

Corrales, J. (2020). Authoritarian survival: Why Maduro hasn't fallen. *Journal of Democracy*, *31*(3), 39–53.

Corrales, J., & Penfold, M. (2011). *Dragon in the tropics: Hugo Chávez and the political economy of revolution in Venezuela*. Brookings Institution Press.

Davenport, C. (2007). State repression and political order. *Annual Review of Political Science*, *10*(1), 1–23.

Di Lonardo, L., Sun, J. S., & Tyson, S. A. (2026). Repression and autocratic consolidation [Forthcoming]. *Games and Economic Behavior*.

Diamond, L. (2002). Elections without democracy: Thinking about hybrid regimes. *Journal of Democracy*, *13*(2), 21–35.

Dragu, T., & Przeworski, A. (2019). Preventive repression: Two types of moral hazard. *American Political Science Review*, *113*(1), 77–87.

Edmond, C. (2013). Information manipulation, coordination, and regime change. *The Review of Economic Studies*, *80*(4), 1422–1458.

Egorov, G., & Sonin, K. (2024). The political economics of non-democracy. *Journal of Economic Literature*, *62*(2), 594–636.

Esberg, J. (2021). Anticipating dissent: The repression of politicians in Pinochet's Chile. *The Journal of Politics*, *83*(2), 689–705.

Escribà-Folch, A. (2013). Repression, political threats, and survival under autocracy. *International Political Science Review*, *34*(5), 543–560.

Esen, B., & Gümüşçü, Ş. (2016). Rising competitive authoritarianism in Turkey. *Third World Quarterly*, *37*(9), 1581–1606.

Esen, B., & Gümüşçü, Ş. (2019). Killing competitive authoritarianism softly: The 2019 local elections in Turkey. *South European Society and Politics*, *24*(3), 317–342.

Fearon, J. D. (2011). Self-enforcing democracy. *The Quarterly Journal of Economics*, *126*(4), 1661–1708.

Fein, H. (1995). Life-integrity violations and democracy in the world, 1987. *Human Rights Quarterly*, *17*(1), 170–191.

Gamboa, L. (2022). *Resisting backsliding: Opposition strategies against the erosion of democracy*. Cambridge University Press.

Gandhi, J. (2015). Elections and political regimes. *Government and Opposition*, *50*(3), 446–468.

Gandhi, J., & Lust-Okar, E. (2009). Elections under authoritarianism. *Annual Review of Political Science*, *12*(1), 403–422.

Gandhi, J., & Przeworski, A. (2007). Authoritarian institutions and the survival of autocrats. *Comparative Political Studies*, *40*(11), 1279–1301.

Gehlbach, S. (2010). Reflections on Putin and the media. *Post-Soviet Affairs*, *26*(1), 77–87.

Gehlbach, S., Sonin, K., & Svolik, M. W. (2016). Formal models of nondemocratic politics. *Annual Review of Political Science*, *19*(1), 565–584.

Gel'man, V. (2013). Cracks in the wall: Challenges to electoral authoritarianism in Russia. *Problems of Post-Communism*, *60*(2), 3–10.

Gel'man, V. (2015). *Authoritarian Russia: Analyzing post-Soviet regime changes*. University of Pittsburgh Press.

Gibilisco, M. (2021). Decentralization, repression, and gambling for unity. *The Journal of Politics*, *83*(4), 1353–1368.

Golosov, G. V. (2011). Russia's regional legislative elections, 2003–2007: Authoritarianism incorporated. *Europe-Asia Studies*, *63*(3), 397–414.

Greene, K. F. (2007). *Why dominant parties lose: Mexico's democratization in comparative perspective*. Cambridge University Press.

Guriev, S., & Treisman, D. (2019). Informational autocrats. *Journal of Economic Perspectives*, *33*(4), 100–127.

Guriev, S., & Treisman, D. (2020). A theory of informational autocracy. *Journal of Public Economics*, *186*, 104158.

Guriev, S., & Treisman, D. (2023). *Spin dictators: The changing face of tyranny in the twenty-first century*. Princeton University Press.

Hager, A., & Krakowski, K. (2022). Does state repression spark protests? Evidence from secret police surveillance in communist Poland. *American Political Science Review*, *116*(2), 564–579.

Hale, H. E. (2014). *Patronal politics: Eurasian regime dynamics in comparative perspective*. Cambridge University Press.

Hassan, M. (2024). Coordinated dis-coordination. *American Political Science Review*, *118*(1), 163–177.

Helms, L. (2023). Political oppositions in democratic and authoritarian regimes: A state-of-the-field(s) review. *Government and Opposition*, *58*(2), 391–414.

Hollyer, J. R., Rosendorff, B. P., & Vreeland, J. R. (2015). Transparency, protest, and autocratic instability. *American Political Science Review*, *109*(4), 764–784.

Howard, M. M., & Roessler, P. G. (2006). Liberalizing electoral outcomes in competitive authoritarian regimes. *American Journal of Political Science*, *50*(2), 365–381.

Hyde, S. D., & Marinov, N. (2012). Which elections can be lost? *Political Analysis*, *20*(2), 191–210.

Kassam, A., & Garamvolgyi, F. (2026). Péter Magyar sworn in as Hungary's prime minister to end 16-year Orbán era. *The Guardian*. Retrieved August 1, 2026, from https://www.theguardian.com/world/2026/may/09/hungary-prime-minister-peter-magyar-sworn-in-viktor-oban

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, *107*(2), 326–343.

Krekó, P., & Enyedi, Z. (2018). Explaining Eastern Europe: Orbán's laboratory of illiberalism. *Journal of Democracy*, *29*(3), 39–51.

Lagunoff, R. (2025). A dynamic model of authoritarian social control. *The Review of Economic Studies*, *92*(5), 3208–3244.

LeBas, A. (2011). *From protest to parties: Party-building and democratization in Africa*. Oxford University Press.

LeBas, A., & Young, L. E. (2024). Repression and dissent in moments of uncertainty: Panel data evidence from Zimbabwe. *American Political Science Review*, *118*(2), 584–601.

Levitsky, S., & Loxton, J. (2013). Populism and competitive authoritarianism in the Andes. *Democratization*, *20*(1), 107–136.

Levitsky, S., & Way, L. A. (2010). *Competitive authoritarianism: Hybrid regimes after the Cold War*. Cambridge University Press.

Lichbach, M. I. (1995). *The rebel's dilemma*. University of Michigan Press.

Lorentzen, P. (2014). China's strategic censorship. *American Journal of Political Science*, *58*(2), 402–414.

Magaloni, B. (2006). *Voting for autocracy: Hegemonic party survival and its demise in Mexico*. Cambridge University Press.

Meng, A. (2021). Ruling parties in authoritarian regimes: Rethinking institutional strength. *British Journal of Political Science*, *51*(2), 526–540.

Mitchell, A. M., Inata, K., & Higashijima, M. (2026). Political and economic protests in authoritarian regimes. *Public Choice*, *206*(1), 283–307.

Montagnes, B. P., & Wolton, S. (2019). Mass purges: Top-down accountability in autocracy. *American Political Science Review*, *113*(4), 1045–1059.

Moore, W. H. (2000). The repression of dissent: A substitution model of government coercion. *Journal of Conflict Resolution*, *44*(1), 107–127.

Morris, S., & Shadmehr, M. (2024). Repression and repertoires. *American Economic Review: Insights*, *6*(3), 413–433.

Olson, M. (1965). *The logic of collective action: Public goods and the theory of groups*. Harvard University Press.

OSCE Office for Democratic Institutions and Human Rights. (2026). *Hungary, parliamentary elections, 12 April 2026: Statement of preliminary findings and conclusions*. Organization for Security and Co-operation in Europe. Retrieved August 1, 2026, from https://odihr.osce.org/odihr/663241

Pierskalla, J. H. (2010). Protest, deterrence, and escalation: The strategic calculus of government repression. *Journal of Conflict Resolution*, *54*(1), 117–145.

Poe, S. C., & Tate, C. N. (1994). Repression of human rights to personal integrity in the 1980s: A global analysis. *American Political Science Review*, *88*(4), 853–872.

Przeworski, A. (2023). Formal models of authoritarian regimes: A critique. *Perspectives on Politics*, *21*(3), 979–988.

Reuter, O. J. (2017). *The origins of dominant parties: Building authoritarian institutions in post-Soviet Russia*. Cambridge University Press.

Ritter, E. H., & Conrad, C. R. (2016). Preventing and responding to dissent: The observational challenges of explaining strategic repression. *American Political Science Review*, *110*(1), 85–99.

Roberts, M. E. (2018). *Censored: Distraction and diversion inside China's Great Firewall*. Princeton University Press.

Robertson, G. B. (2011). *The politics of protest in hybrid regimes: Managing dissent in post-communist Russia*. Cambridge University Press.

Rosenfeld, B., & Wallace, J. (2024). Information politics and propaganda in authoritarian societies. *Annual Review of Political Science*, *27*(1), 263–281.

Roy, A., & MacDonald, L. (2026). The interplay between discrimination, repression and inequality under authoritarianism. *European Journal of Political Economy*, *95*, 102877.

Rozenas, A. (2020). A theory of demographically targeted repression. *Journal of Conflict Resolution*, *64*(7–8), 1254–1278.

Rozenas, A., & Zhukov, Y. M. (2019). Mass repression and political loyalty: Evidence from Stalin's terror by hunger. *American Political Science Review*, *113*(2), 569–583.

Schedler, A. (2013). *The politics of uncertainty: Sustaining and subverting electoral authoritarianism*. Oxford University Press.

Scheppele, K. L. (2022). How Viktor Orbán wins. *Journal of Democracy*, *33*(3), 45–61.

Shadmehr, M., & Bernhardt, D. (2011). Collective action with uncertain payoffs: Coordination, public signals, and punishment dilemmas. *American Political Science Review*, *105*(4), 829–851.

Siegel, D. A. (2011). When does repression work? Collective action in social networks. *The Journal of Politics*, *73*(4), 993–1010.

Slantchev, B. L., & Matush, K. S. (2020). The authoritarian wager: Political action and the sudden collapse of repression. *Comparative Political Studies*, *53*(2), 214–252.

Snyder, J. M., & Ting, M. M. (2002). An informational rationale for political parties. *American Journal of Political Science*, *46*(1), 90–110.

Somer, M. (2016). Understanding Turkey's democratic breakdown: Old vs. new and indigenous vs. global authoritarianism. *Southeast European and Black Sea Studies*, *16*(4), 481–503.

Sullivan, C. M., & Davenport, C. (2017). The rebel alliance strikes back: Understanding the politics of backlash mobilization. *Mobilization*, *22*(1), 39–56.

Svolik, M. W. (2012). *The politics of authoritarian rule*. Cambridge University Press.

Szeidl, A., & Szűcs, F. (2021). Media capture through favor exchange. *Econometrica*, *89*(1), 281–310.

Treisman, D. (2011). *The return: Russia's journey from Gorbachev to Medvedev*. Free Press.

Tucker, J. A. (2007). Enough! Electoral fraud, collective action problems, and post-communist colored revolutions. *Perspectives on Politics*, *5*(3), 535–551.

Tullock, G. (1980). Efficient rent seeking. In J. M. Buchanan, R. D. Tollison, & G. Tullock (Eds.), *Toward a theory of the rent-seeking society* (pp. 97–112). Texas A&M University Press.

Tyson, S. A. (2018). The agency problem underlying repression. *The Journal of Politics*, *80*(4), 1297–1310.

Wahman, M. (2013). Opposition coalitions and democratization by election. *Government and Opposition*, *48*(1), 3–32.

White, D., & Herzog, M. (2016). Examining state capacity in the context of electoral authoritarianism, regime formation and consolidation in Russia and Turkey. *Southeast European and Black Sea Studies*, *16*(4), 551–569.

Wintrobe, R. (1990). The tinpot and the totalitarian: An economic theory of dictatorship. *American Political Science Review*, *84*(3), 849–872.

Wright, J., & Escribà-Folch, A. (2012). Authoritarian institutions and regime survival: Transitions to democracy and subsequent autocracy. *British Journal of Political Science*, *42*(2), 283–309.

Xu, X. (2021). To repress or to co-opt? Authoritarian control in the age of digital surveillance. *American Journal of Political Science*, *65*(2), 309–325.

Yavuz, M. H., & Koç, R. (2024). Do Turkey's 2024 local elections signal the end of Erdoğan's reign? *Middle East Policy*, *31*(2), 95–107.

Young, L. E. (2019). The psychology of state repression: Fear and dissent decisions in Zimbabwe. *American Political Science Review*, *113*(1), 140–155.

A Proofs and formalities

A.1 Formal analysis of the defection equilibrium

Lemma A.1. ***Comparative statics of*** $V_i^D(n)$
The inactivity value $V_i^D(n)$ *satisfies the following comparative statics:*

A.1.1 $V_i^D(n)$ *is strictly increasing and strictly concave in* n*.*

A.1.2 $V_i^D(n)$ *is strictly increasing in* Π_0*.*

A.1.3 $V_i^D(n)$ *is strictly increasing in* β*.*

A.1.4 $V_i^D(n)$ *is strictly decreasing in* ζ_d*.*

Proof: Recall the closed form

$$V_i^D(n) = \frac{n\Pi_0}{\zeta_d(1-\beta)+n}.$$

Throughout, the denominator $\zeta_d(1-\beta)+n$ is strictly positive, since $\zeta_d > 0$, $\beta \in (0,1)$, and $n \geq 1$.

A.1.1: Differentiating with respect to n gives

$$\frac{\partial V_i^D(n)}{\partial n} = \frac{\zeta_d(1-\beta)\Pi_0}{\Big(\zeta_d(1-\beta)+n\Big)^2} > 0,$$

and

$$\frac{\partial^2 V_i^D(n)}{\partial n^2} = -2\frac{\zeta_d(1-\beta)\Pi_0}{\Big(\zeta_d(1-\beta)+n\Big)^3} < 0.$$

Hence $V_i^D(n)$ is strictly increasing and strictly concave in n.

A.1.2: Differentiating with respect to Π_0 gives

$$\frac{\partial V_i^D(n)}{\partial \Pi_0} = \frac{n}{\zeta_d(1-\beta)+n} > 0.$$

Hence $V_i^D(n)$ is strictly increasing in Π_0.

A.1.3: Differentiating with respect to β gives

$$\frac{\partial V_i^D(n)}{\partial \beta} = \frac{n\Pi_0\zeta_d}{\Big(\zeta_d(1-\beta)+n\Big)^2} > 0.$$

Hence $V_i^D(n)$ is strictly increasing in β.

A.1.4: Differentiating with respect to ζ_d gives

$$\frac{\partial V_i^D(n)}{\partial \zeta_d} = -\frac{n\Pi_0(1-\beta)}{\Big(\zeta_d(1-\beta)+n\Big)^2} < 0.$$

Hence $V_i^D(n)$ is strictly decreasing in ζ_d. □

Lemma A.2. *Comparative statics of* $V_d^D(n)$
The activity value $V_d^D(n)$ *satisfies the following comparative statics:*

A.2.1 $V_d^D(n)$ *is strictly increasing in* Π_1.

A.2.2 $V_d^D(n)$ *is increasing in* Π_0, *and strictly increasing in* Π_0 *whenever* $n > 1$.

A.2.3 $V_d^D(n)$ *is strictly decreasing in* R.

A.2.4 If $V_d^D(n) > 0$, *then* $V_d^D(n)$ *is strictly increasing in* β.

A.2.5 $V_d^D(n)$ *is strictly decreasing in* ζ_d.

Proof: Since $\Delta\Pi = \Pi_1 - \Pi_0$, the closed form can be written as

$$V_d^D(n) = \frac{\Delta\Pi + n\Pi_0 - R(\zeta_d + n)}{\zeta_d(1-\beta) + n} = \frac{\Pi_1 + (n-1)\Pi_0 - R(\zeta_d + n)}{\zeta_d(1-\beta) + n}.$$

Throughout, the denominator $\zeta_d(1-\beta) + n$ is strictly positive, since $\zeta_d > 0$, $\beta \in (0, 1)$, and $n \geq 1$.

A.2.1: Differentiating with respect to Π_1 gives

$$\frac{\partial V_d^D(n)}{\partial \Pi_1} = \frac{1}{\zeta_d(1-\beta) + n} > 0.$$

Hence $V_d^D(n)$ is strictly increasing in Π_1.

A.2.2: Differentiating with respect to Π_0 gives

$$\frac{\partial V_d^D(n)}{\partial \Pi_0} = \frac{n-1}{\zeta_d(1-\beta) + n} \geq 0.$$

The inequality is strict whenever $n > 1$. Hence $V_d^D(n)$ is increasing in Π_0, and strictly increasing in Π_0 whenever $n > 1$.

A.2.3: Differentiating with respect to R gives

$$\frac{\partial V_d^D(n)}{\partial R} = -\frac{\zeta_d + n}{\zeta_d(1-\beta) + n} < 0.$$

Hence $V_d^D(n)$ is strictly decreasing in R.

A.2.4: Differentiating with respect to β gives

$$\frac{\partial V_d^D(n)}{\partial \beta} = \frac{\zeta_d\Big(\Delta\Pi + n\Pi_0 - R(\zeta_d + n)\Big)}{\Big(\zeta_d(1-\beta)+n\Big)^2} = \frac{\zeta_d V_d^D(n)}{\zeta_d(1-\beta)+n}.$$

This expression is strictly positive whenever $V_d^D(n) > 0$. Hence $V_d^D(n)$ is strictly increasing in β whenever $V_d^D(n) > 0$.

A.2.5: Differentiating with respect to ζ_d gives

$$\frac{\partial V_d^D(n)}{\partial \zeta_d} = -\frac{\beta R n + (1-\beta)(\Delta\Pi + n\Pi_0)}{\Big(\zeta_d(1-\beta)+n\Big)^2} < 0.$$

Hence $V_d^D(n)$ is strictly decreasing in ζ_d. □

Lemma A.3. ***Monotonicity and curvature of*** $V_d^D(n)$ ***in*** n

A.3.1 If

$$\zeta_d\Big((1-\beta)\Pi_0+\beta R\Big) > \Delta\Pi,$$

then $V_d^D(n)$ *is strictly increasing and strictly concave in* n.

A.3.2 If

$$\zeta_d\Big((1-\beta)\Pi_0+\beta R\Big) < \Delta\Pi,$$

then $V_d^D(n)$ *is strictly decreasing and strictly convex in* n.

Proof: Recall the closed form

$$V_d^D(n) = \frac{\Delta\Pi - R\zeta_d + n(\Pi_0 - R)}{\zeta_d(1-\beta)+n}.$$

Differentiating with respect to n gives

$$\frac{\partial V_d^D(n)}{\partial n} = \frac{\zeta_d\Big((1-\beta)\Pi_0+\beta R\Big)-\Delta\Pi}{\Big(\zeta_d(1-\beta)+n\Big)^2},$$

$$\frac{\partial^2 V_d^D(n)}{\partial n^2} = -2\frac{\zeta_d\Big((1-\beta)\Pi_0+\beta R\Big)-\Delta\Pi}{\Big(\zeta_d(1-\beta)+n\Big)^3}.$$

The denominators are strictly positive, since $\zeta_d > 0$, $\beta \in (0,1)$, and $n \geq 1$. Hence the first derivative has the sign of

$$\zeta_d\Big((1-\beta)\Pi_0+\beta R\Big)-\Delta\Pi,$$

while the second derivative has the opposite sign.

A.3.1: If

$$\zeta_d\Big((1-\beta)\Pi_0+\beta R\Big) > \Delta\Pi,$$

then

$$\frac{\partial V_d^D(n)}{\partial n} > 0 \qquad \text{and} \qquad \frac{\partial^2 V_d^D(n)}{\partial n^2} < 0.$$

Hence $V_d^D(n)$ is strictly increasing and strictly concave in n.

A.3.2: If

$$\zeta_d\Big((1-\beta)\Pi_0+\beta R\Big) < \Delta\Pi,$$

then

$$\frac{\partial V_d^D(n)}{\partial n} < 0 \qquad \text{and} \qquad \frac{\partial^2 V_d^D(n)}{\partial n^2} > 0.$$

Hence $V_d^D(n)$ is strictly decreasing and strictly convex in n. $\square$

Lemma A.4. ***Properties of*** $\tilde{n}^D$

A.4.1 The threshold $\tilde{n}^D \geq 0$ *is well defined if and only if*

$$R \leq \frac{\Pi_1}{\zeta_d + 1}.$$

A.4.2 If $\tilde{n}^D \geq 0$ *exists, it is unique.*

A.4.3 If $\tilde{n}^D \geq 0$ *exists, then*

$$\Delta V^{D\prime}(\tilde{n}^D) < 0.$$

A.4.4

$$\tilde{n}^D = \frac{\Delta\Pi - R(\zeta_d(2-\beta)+1) + \sqrt{\begin{array}{l}\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 \\ \quad + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)\end{array}}}{2R}.$$

Proof: Recall that

$$\Delta V^D(n) = V_d^D(n+1) - V_i^D(n),$$

where

$$V_d^D(n+1) = \frac{\Delta\Pi - R\zeta_d + (n+1)(\Pi_0 - R)}{\zeta_d(1-\beta)+n+1}, \qquad V_i^D(n) = \frac{n\Pi_0}{\zeta_d(1-\beta)+n}.$$

The threshold $\tilde{n}^D$ solves $\Delta V^D(\tilde{n}^D) = 0$. Clearing denominators gives $Q(\tilde{n}^D) = 0$, where

$$Q(n) = -Rn^2 + \Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)n + \zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big). \quad (24)$$

A.4.1: The sign of the constant term in (24) is governed by

$$\Pi_1 - R(\zeta_d + 1).$$

If $R \leq \frac{\Pi_1}{\zeta_d+1}$, then the constant term is non-negative. Since the coefficient on

n^2 is $-R < 0$, the parabola opens downward and is non-negative at $n = 0$, and since $Q(n) \to -\infty$ as $n \to +\infty$, it must cross zero for some $n \geq 0$. Hence $\tilde{n}^D \geq 0$ exists. Suppose now $R > \frac{\Pi_1}{\zeta_d+1}$ and real roots exist. The constant term is negative, so Vieta's formulas imply the product of the two roots is positive. Moreover,

$$R > \frac{\Pi_1}{\zeta_d + 1} \implies R(\zeta_d(2-\beta)+1) > R(\zeta_d + 1) > \Pi_1 > \Delta\Pi,$$

so the sum of the two roots,

$$\frac{\Delta\Pi - R(\zeta_d(2-\beta)+1)}{R},$$

is negative. A positive product and a negative sum imply both roots are negative. Hence no non-negative solution exists.

A.4.2: The product of the roots of (24) equals

$$\frac{\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)}{-R},$$

which is non-positive whenever $\tilde{n}^D \geq 0$ exists, by Part A.4.1. Two real roots with non-positive product have weakly opposite signs, so at most one can be non-negative. Hence $\tilde{n}^D \geq 0$, if it exists, is unique.

A.4.3: We first show that if $\tilde{n}^D \geq 0$ exists, then Q has two distinct real roots. The discriminant is

$$\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big).$$

If $R < \frac{\Pi_1}{\zeta_d+1}$, then the second term is strictly positive, so the discriminant is strictly positive. If $R = \frac{\Pi_1}{\zeta_d+1}$, then it equals

$$\Big(-\Pi_0 - R\zeta_d(1-\beta)\Big)^2 > 0.$$

Hence the discriminant is strictly positive, and therefore Q has two distinct real roots.

We can rewrite $\Delta V^D(n)$ as

$$\Delta V^D(n) = \frac{Q(n)}{\Big(\zeta_d(1-\beta)+n+1\Big)\Big(\zeta_d(1-\beta)+n\Big)},$$

with $Q(n)$ the quadratic in (24). Since $Q(\tilde{n}^D) = 0$ and the denominator is strictly positive at $\tilde{n}^D$, it follows that

$$\Delta V^{D\prime}(\tilde{n}^D) = \frac{Q'(\tilde{n}^D)}{\Big(\zeta_d(1-\beta)+\tilde{n}^D+1\Big)\Big(\zeta_d(1-\beta)+\tilde{n}^D\Big)}.$$

Since $Q(n)$ is a downward-opening parabola and $\tilde{n}^D$ is the larger of its two real roots, we have $Q'(\tilde{n}^D) < 0$ and thus $\Delta V^{D\prime}(\tilde{n}^D) < 0$.

A.4.4: The threshold $\tilde{n}^D$ solves (24). Treating it as a quadratic in n and applying the quadratic formula, the two roots are

$$\begin{aligned} n &= \frac{-\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big) \pm \sqrt{\begin{aligned}&\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 \\ &\quad + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)\end{aligned}}}{-2R} \\ &= \frac{\Delta\Pi - R(\zeta_d(2-\beta)+1) \mp \sqrt{\begin{aligned}&\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 \\ &\quad + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)\end{aligned}}}{2R}. \end{aligned}$$

The non-negative defection-entry threshold must be the larger of the two, and thus the root taking the positive square root,

$$\tilde{n}^D = \frac{\Delta\Pi - R(\zeta_d(2-\beta)+1) + \sqrt{\begin{aligned}&\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 \\ &\quad + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)\end{aligned}}}{2R}.$$

□

Proof of Lemma 1

Consider any $\sigma \in D_x$ such that $n(\sigma, \emptyset) = n^D$. Moreover, consider any $\hat{\sigma} \in D_x$ such that $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ for all $\tilde{f} \neq f$. Let $\hat{\sigma}_{\tilde{f}} = (\hat{x}_{\tilde{f}}, \hat{y}_{\tilde{f}})$ and $\sigma_{\tilde{f}} = (x_{\tilde{f}}, y_{\tilde{f}})$ for any $\tilde{f}$.

Case 1: $\hat{x}_f = x_f$. Then no faction changes its participation decision under $\hat{\sigma}$ as compared to σ, that is, $\{x_{\tilde{f}}\}_{\tilde{f}\in F} = \{\hat{x}_{\tilde{f}}\}_{\tilde{f}\in F}$, and thus $n(\sigma, \emptyset) = n(\hat{\sigma}, \emptyset)$. Since $\hat{\sigma}, \sigma \in D_x$, $m^t(\sigma, \emptyset) = m^t(\hat{\sigma}, \emptyset) = 0$ for all $t \geq 1$. All active factions always defect along the path of play under $\hat{\sigma}$ and σ. Since no faction changes its on-path play at $\hat{\sigma}$ as compared to σ, the expected present values of f coincide: $U_f^0(\sigma, \emptyset) = U_f^0(\hat{\sigma}, \emptyset)$.

Case 2: $x_f = a$ *and* $\hat{x}_f = i$. Faction f being inactive under $\hat{\sigma}$ and active under σ implies $n(\hat{\sigma}, \emptyset) = n^D - 1$. Moreover, since $\hat{\sigma}, \sigma \in D_x$, $m^t(\sigma, \emptyset) = m^t(\hat{\sigma}, \emptyset) = 0$ for all $t \geq 1$. Thus, the resulting expected present values are $U_f^0(\sigma, \emptyset) = V_d^D(n^D)$ and $U_f^0(\hat{\sigma}, \emptyset) = V_i^D(n^D - 1)$. By definition of $\tilde{n}^D$, $V_d^D(\tilde{n}^D + 1) - V_i^D(\tilde{n}^D) = 0$. By Lemma A.4, $V_d^D(x) - V_i^D(x-1)$ is decreasing in x, and by definition of n^D, $n^D \leq \tilde{n}^D + 1$. Hence, $V_d^D(n^D) - V_i^D(n^D - 1) \geq 0$ and thus $U_f^0(\sigma, \emptyset) \geq U_f^0(\hat{\sigma}, \emptyset)$.

Case 3: $x_f = i$ *and* $\hat{x}_f = a$. Faction f being active under $\hat{\sigma}$ and inactive under σ implies $n(\hat{\sigma}, \emptyset) = n^D + 1$. Moreover, since $\hat{\sigma}, \sigma \in D_x$, $m^t(\sigma, \emptyset) = m^t(\hat{\sigma}, \emptyset) = 0$ for all $t \geq 1$. Thus, the resulting expected present values are $U_f^0(\sigma, \emptyset) = V_i^D(n^D)$ and $U_f^0(\hat{\sigma}, \emptyset) = V_d^D(n^D + 1)$. By definition of $\tilde{n}^D$, $V_d^D(\tilde{n}^D + 1) - V_i^D(\tilde{n}^D) = 0$. By Lemma A.4, $V_d^D(x+1) - V_i^D(x)$ is decreasing in x, and by definition of n^D, $n^D \geq \tilde{n}^D$. Hence, $V_d^D(n^D+1) - V_i^D(n^D) \leq 0$ and thus $U_f^0(\sigma, \emptyset) \geq U_f^0(\hat{\sigma}, \emptyset)$. $\square$

Lemma A.5. ***Comparative statics of*** $\tilde{n}^D$
Suppose that $\tilde{n}^D > 0$ *exists. Then the defection-entry threshold satisfies the following comparative statics:*

A.5.1 $\tilde{n}^D$ *is strictly increasing in* Π_1*.*

A.5.2 $\tilde{n}^D$ *is strictly decreasing in* Π_0*.*

A.5.3 $\tilde{n}^D$ *is strictly decreasing in* R*.*

A.5.4 $\tilde{n}^D$ *is strictly decreasing in* β*.*

Proof: Recall that

$$V_d^D(n+1) = \frac{\Delta\Pi - R\zeta_d + (n+1)(\Pi_0 - R)}{\zeta_d(1-\beta)+n+1}, \qquad V_i^D(n) = \frac{n\Pi_0}{\zeta_d(1-\beta)+n}.$$

The defection-entry threshold $\tilde{n}^D$ solves

$$\Delta V^D(\tilde{n}^D) = V_d^D(\tilde{n}^D+1) - V_i^D(\tilde{n}^D) = 0.$$

For any scalar parameter θ, the implicit function theorem gives

$$\frac{\partial \tilde{n}^D}{\partial \theta} = -\frac{\partial \Delta V^D/\partial \theta}{\partial \Delta V^D/\partial \tilde{n}^D}.$$

By Lemma A.4,

$$\frac{\partial \Delta V^D}{\partial \tilde{n}^D} < 0$$

at the equilibrium intersection. Hence the sign of $\partial \tilde{n}^D/\partial\theta$ is the same as the sign of $\partial \Delta V^D/\partial\theta$. The denominators $\zeta_d(1-\beta)+\tilde{n}^D$ and $\zeta_d(1-\beta)+\tilde{n}^D+1$ are strictly positive, since $\zeta_d > 0$, $\beta \in (0,1)$, and $\tilde{n}^D > 0$.

A.5.1: Differentiating ΔV^D with respect to Π_1 gives

$$\frac{\partial \Delta V^D}{\partial \Pi_1} = \frac{1}{\zeta_d(1-\beta)+\tilde{n}^D+1} > 0.$$

Therefore,

$$\frac{\partial \tilde{n}^D}{\partial \Pi_1} > 0.$$

A.5.2: Differentiating ΔV^D with respect to Π_0 gives

$$\begin{aligned}\frac{\partial \Delta V^D}{\partial \Pi_0} &= \frac{\tilde{n}^D}{\zeta_d(1-\beta)+\tilde{n}^D+1} - \frac{\tilde{n}^D}{\zeta_d(1-\beta)+\tilde{n}^D} \\ &= -\frac{\tilde{n}^D}{\Big(\zeta_d(1-\beta)+\tilde{n}^D\Big)\Big(\zeta_d(1-\beta)+\tilde{n}^D+1\Big)} < 0.\end{aligned}$$

Therefore,

$$\frac{\partial \tilde{n}^D}{\partial \Pi_0} < 0.$$

A.5.3: Differentiating ΔV^D with respect to R gives

$$\frac{\partial \Delta V^D}{\partial R} = -\frac{\zeta_d+\tilde{n}^D+1}{\zeta_d(1-\beta)+\tilde{n}^D+1} < 0.$$

Therefore,

$$\frac{\partial \tilde{n}^D}{\partial R} < 0.$$

A.5.4: Differentiating the two value functions with respect to β gives

$$\frac{\partial V_i^D(\tilde{n}^D)}{\partial \beta} = \zeta_d \frac{V_i^D(\tilde{n}^D)}{\zeta_d(1-\beta)+\tilde{n}^D}, \qquad \frac{\partial V_d^D(\tilde{n}^D+1)}{\partial \beta} = \zeta_d \frac{V_d^D(\tilde{n}^D+1)}{\zeta_d(1-\beta)+\tilde{n}^D+1}.$$

At the threshold,

$$V_d^D(\tilde{n}^D+1) = V_i^D(\tilde{n}^D).$$

Hence

$$\begin{aligned}\frac{\partial \Delta V^D}{\partial \beta} &= \zeta_d V_i^D(\tilde{n}^D)\left(\frac{1}{\zeta_d(1-\beta)+\tilde{n}^D+1} - \frac{1}{\zeta_d(1-\beta)+\tilde{n}^D}\right) \\ &= -\frac{\zeta_d V_i^D(\tilde{n}^D)}{\Big(\zeta_d(1-\beta)+\tilde{n}^D\Big)\Big(\zeta_d(1-\beta)+\tilde{n}^D+1\Big)} < 0,\end{aligned}$$

where $V_i^D(\tilde{n}^D) > 0$ since $\Pi_0 > 0$ and $\tilde{n}^D > 0$. Therefore,

$$\frac{\partial \tilde{n}^D}{\partial \beta} < 0.$$

$\square$

Lemma A.6. *Comparative statics of* $\tilde{n}^D$ *in* ζ_d
Suppose that $\tilde{n}^D > 0$ *exists. Then*

$$\frac{\partial \tilde{n}^D}{\partial \zeta_d} \le 0 \quad \Longleftrightarrow \quad \zeta_d \ge \kappa,$$

where $\kappa \in [0, \infty)$ *is defined by*

$$\kappa := \begin{cases} 0 & \text{if } \Pi_1 > R + (2-\beta)\Pi_0, \\ \dfrac{\Pi_1 - R - \dfrac{2-\beta}{\beta}\Pi_0 \left(1 - \sqrt{\dfrac{\Pi_0 + \beta R - \beta \Pi_1}{\Pi_0}}\right)}{\beta R} & \text{otherwise.} \end{cases}$$

Proof: Recall the closed-form solution of $\tilde{n}^D$ as defined in Lemma A.4. Differentiating it with respect to ζ_d gives

$$\frac{\partial \tilde{n}^D}{\partial \zeta_d} = -\frac{2-\beta}{2} + \frac{R\beta^2\zeta_d + R\beta + (2-\beta)\Pi_0 - \beta\Pi_1}{2\sqrt{\begin{aligned}&\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 \\ &\quad + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)\end{aligned}}}. \tag{25}$$

Since $\tilde{n}^D > 0$ exists, Lemma A.4 implies

$$R < \frac{\Pi_1}{\zeta_d + 1},$$

so the square-root term in (25) is strictly positive. Therefore,

$$\begin{aligned} \frac{\partial \tilde{n}^D}{\partial \zeta_d} \le 0 \quad \Longleftrightarrow \quad & R\beta^2\zeta_d + R\beta + (2-\beta)\Pi_0 - \beta\Pi_1 \\ & \le (2-\beta)\sqrt{\begin{aligned}&\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 \\ &\quad + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)\end{aligned}}. \end{aligned} \tag{26}$$

If

$$R\beta^2\zeta_d + R\beta + (2-\beta)\Pi_0 - \beta\Pi_1 \le 0,$$

then (26) holds automatically, since its right-hand side is non-negative. This is the case summarized by $\kappa = 0$.

Now suppose instead that

$$R\beta^2\zeta_d + R\beta + (2-\beta)\Pi_0 - \beta\Pi_1 > 0.$$

Then both sides of (26) are non-negative, so squaring preserves the inequality, which yields

$$\begin{aligned}(2-\beta)^2 &\left[\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)\right] \\ &\geq \Big(R\beta^2\zeta_d + R\beta + (2-\beta)\Pi_0 - \beta\Pi_1\Big)^2.\end{aligned}$$

Expanding the bracket on the left-hand side as a polynomial in ζ_d,

$$\begin{aligned}&\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big) \\ &= R^2(2-\beta)^2\zeta_d^2 - 2R(2-\beta)\Big(\Delta\Pi - R\Big)\zeta_d + \Big(\Delta\Pi - R\Big)^2 \\ &\quad + 4R(1-\beta)\Pi_1\zeta_d - 4R^2(1-\beta)\zeta_d^2 - 4R^2(1-\beta)\zeta_d \\ &= R^2\beta^2\zeta_d^2 + 2R\Big((2-\beta)(R-\Delta\Pi) + 2(1-\beta)\Pi_1 - 2R(1-\beta)\Big)\zeta_d + \Big(\Delta\Pi - R\Big)^2,\end{aligned}$$

where the coefficient on ζ_d^2 collapses because

$$R^2(2-\beta)^2 - 4R^2(1-\beta) = R^2\Big(4 - 4\beta + \beta^2 - 4 + 4\beta\Big) = R^2\beta^2.$$

Multiplying this polynomial by $(2-\beta)^2$ and subtracting the square on the right-hand side, the inequality becomes

$$\begin{aligned}&\Big((2-\beta)^2R^2\beta^2 - R^2\beta^4\Big)\zeta_d^2 \\ &\quad + \Big(2(2-\beta)^2R\Big((2-\beta)(R-\Delta\Pi) + 2(1-\beta)\Pi_1 - 2R(1-\beta)\Big) \\ &\qquad\qquad - 2R\beta^2\Big(R\beta + (2-\beta)\Pi_0 - \beta\Pi_1\Big)\Big)\zeta_d \\ &\quad + \Big((2-\beta)^2\Big(\Delta\Pi - R\Big)^2 - \Big(R\beta + (2-\beta)\Pi_0 - \beta\Pi_1\Big)^2\Big) \geq 0.\end{aligned}$$

The coefficient on ζ_d^2 simplifies to

$$(2-\beta)^2R^2\beta^2 - R^2\beta^4 = R^2\beta^2\Big((2-\beta)^2 - \beta^2\Big) = 4R^2\beta^2(1-\beta),$$

which is strictly positive, since $R > 0$ and $\beta \in (0,1)$, so the parabola in ζ_d

opens upward. Its two roots are

$$\zeta_d = \frac{\Pi_1 - R - \dfrac{2-\beta}{\beta}\Pi_0 \left(1 \mp \sqrt{\dfrac{\Pi_0 + \beta R - \beta \Pi_1}{\Pi_0}}\right)}{\beta R},$$

and on the region where $\tilde{n}^D > 0$ exists only the larger root, obtained with the minus sign inside the parenthesis, lies in the admissible range. The inequality is therefore equivalent to

$$\zeta_d \geq \frac{\Pi_1 - R - \frac{2-\beta}{\beta}\Pi_0 \left(1 - \sqrt{\frac{\Pi_0 + \beta R - \beta \Pi_1}{\Pi_0}}\right)}{\beta R}.$$

Thus, in this case,

$$\frac{\partial \tilde{n}^D}{\partial \zeta_d} \leq 0 \quad \Longleftrightarrow \quad \zeta_d \geq \kappa.$$

Combining the two cases gives

$$\frac{\partial \tilde{n}^D}{\partial \zeta_d} \leq 0 \quad \Longleftrightarrow \quad \zeta_d \geq \kappa.$$

□

Lemma A.7. ***Comparative statics of*** n^D

A.7.1 n^D *is increasing in* Π_1.

A.7.2 n^D *is decreasing in* Π_0.

A.7.3 n^D *is decreasing in* R.

A.7.4 n^D *is decreasing in* β.

A.7.5 n^D *is decreasing in* ζ_d *if* $\zeta_d \geq \kappa$ *and increasing in* ζ_d *if* $\zeta_d \leq \kappa$, *where* κ *is defined as in Lemma A.6.*

Proof: Recall that n^D is the smallest integer larger than $\tilde{n}^D$, so n^D is a weakly monotonic transformation of $\tilde{n}^D$, changing value only when $\tilde{n}^D$ crosses an integer threshold and remaining constant in between. Hence, whenever $\tilde{n}^D$ is strictly increasing (decreasing) in some variable θ, n^D is increasing (decreasing) in θ, with jumps occurring only where $\tilde{n}^D$ takes an integer value. The lemma then immediately follows from Lemma A.5 and Lemma A.6. $\square$

Lemma A.8. ***Repression and*** $\tilde{n}^D$

Suppose that $\tilde{n}^D > 0$ *exists. Then* $\tilde{n}^D$ *satisfies the following properties:*

A.8.1 $\lim_{R\to 0^+} \tilde{n}^D = \infty$.

A.8.2 $\lim_{R\to \Pi_1/(\zeta_d+1)^-} \tilde{n}^D = 0$.

Proof: Recall the closed-form solution of $\tilde{n}^D$ as defined in Lemma A.4:

$$\tilde{n}^D = \frac{\Delta\Pi - R(\zeta_d(2-\beta)+1) + \sqrt{\begin{aligned}&\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2\\ &\quad + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)\end{aligned}}}{2R}.$$

A.8.1: As $R \to 0$, the numerator stays bounded and $\tilde{n}^D$ exists since $0 < \frac{\Pi_1}{\zeta_d+1}$. However, the denominator $2R \to 0$ and thus $\tilde{n}^D \to \infty$.

A.8.2: Substituting the limit $R = \frac{\Pi_1}{\zeta_d+1}$ into $\tilde{n}^D$ yields

$$\tilde{n}^D = \frac{\Delta\Pi - \frac{\Pi_1}{\zeta_d+1}(\zeta_d(2-\beta)+1) + \sqrt{(\Pi_1 - \Pi_0 - \frac{\Pi_1}{\zeta_d+1}(\zeta_d(2-\beta)+1))^2}}{2\frac{\Pi_1}{\zeta_d+1}}.$$

Note that $\Delta\Pi - \frac{\Pi_1}{\zeta_d+1}(\zeta_d(2-\beta)+1) < 0$ since $\Delta\Pi = \Pi_1 - \Pi_0$ and $\zeta_d + 1 < (\zeta_d(2-\beta)+1)$. Thus, the numerator is 0 and hence $\tilde{n}^D = 0$ at $R = \frac{\Pi_1}{\zeta_d+1}$. $\square$

Proof of Lemma 2

First, note that by definition of n^D, $n^D \geq \tilde{n}^D$ for all $R \in (0, \frac{\Pi_1}{\zeta_d+1}]$. Hence, $\lim_{R\to 0^+} n^D \geq \lim_{R\to 0^+} \tilde{n}^D$. From Lemma A.8.1, we know that $\lim_{R\to 0^+} \tilde{n}^D = \infty$. It follows that $\lim_{R\to 0^+} n^D = \infty$. Second, recall from Lemma A.8.2 that $\lim_{R\to \Pi_1/(\zeta_d+1)^-} \tilde{n}^D = 0$. Hence, there is some $\epsilon > 0$ such that $\tilde{n}^D \in [0, 1)$ for all $R \in (\frac{\Pi_1}{\zeta_d+1} - \epsilon, \frac{\Pi_1}{\zeta_d+1}]$. $\tilde{n}^D \in [0, 1)$ implies $n^D = 1$, and thus $n^D = 1$ for all $R \in (\frac{\Pi_1}{\zeta_d+1} - \epsilon, \frac{\Pi_1}{\zeta_d+1}]$. Hence, $\lim_{R\to \Pi_1/(\zeta_d+1)^-} n^D = 1$. Third, the last statement follows directly from Lemma A.7. $\square$

Lemma A.9. ***A sole active faction cooperates***
Consider any subgame perfect Nash equilibrium σ. For any $t \geq 1$ and $h \in H^t$:

$$n(\sigma, h) = 1 \Rightarrow m^t(\sigma, h) = 1.$$

Proof: Consider any subgame perfect Nash equilibrium σ, any $t \geq 1$, $h \in H^t$ with $n(\sigma, h) = 1$, and let f be the sole element of $F_a(\sigma, h)$. Since participation is fixed at $t = 0$, $n(\sigma, \tilde{h}) = 1$ at every successor history $\tilde{h}$ of h, so f faces no strategic interaction from h onward, and simply maximizes at every such history.

Let $\tilde{p}(c) = p(1, 1)$ and $\tilde{p}(d) = p(1, 0)$. The decision problem f faces is identical at h and every successor history $\tilde{h}$ of h, independent of how h and $\tilde{h}$ were reached. Since σ is a subgame perfect Nash equilibrium, it prescribes an optimal action at every such history. Because the continuation problem is identical at every such history, the same action $y^* \in \{c, d\}$ must be optimal at each of them. The corresponding continuation value satisfies

$$V^* = -R + \Big(1 - \tilde{p}(y^*)\Big)\Pi_1 + \tilde{p}(y^*)\,\beta\,V^* \quad \Longrightarrow \quad V^* = \frac{-R + \Big(1 - \tilde{p}(y^*)\Big)\Pi_1}{1 - \tilde{p}(y^*)\,\beta}.$$

Let $V(p) := \dfrac{-R + (1-p)\Pi_1}{1 - p\beta}$ for $p \in [0, 1)$, so that $V^* = \max\{V(\tilde{p}(c)), V(\tilde{p}(d))\}$. Since

$$V'(p) = -\frac{\Pi_1(1-\beta) + \beta R}{(1 - p\beta)^2} < 0$$

for all $p \in [0, 1)$ and $\tilde{p}(c) = p(1, 1) < p(1, 0) = \tilde{p}(d)$, it follows that $V(\tilde{p}(c)) > V(\tilde{p}(d))$. Thus, $y^* = c$ is the unique maximizer.

Hence σ prescribes c at h and at every successor history $\tilde{h}$ of h. Since σ, t, and h were arbitrary, the lemma follows. $\square$

Proof of Lemma 3

Consider any σ as described, and let $\sigma_f = (x_f, y_f)$ for all $f \in F$. Note that the two conditions on σ are thus equivalent to

1. $n(\sigma, h) \geq 2 \Rightarrow y_f(h) = d$ and

2. $n(\sigma, h) = 1 \Rightarrow y_f(h) = c$

for all $f \in F_a(\sigma, h)$ and all $t \geq 1$, $h \in H^t$. In the following, consider any $f \in F$, $t \geq 1$, $h \in H^t$, and $\hat{\sigma} \neq \sigma$ such that

1. $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ for all $\tilde{f} \neq f$ and

2. $\hat{\sigma}_f(\tilde{h}) = \sigma_f(\tilde{h})$ for all $\tilde{h} \neq h$.

Let $\hat{\sigma}_{\tilde{f}} = (\hat{x}_{\tilde{f}}, \hat{y}_{\tilde{f}})$ for all $\tilde{f} \in F$. If $f \notin F_a(\sigma, h)$, then since participation is fixed at $t = 0$, f takes no actions at h or any future history. Thus $U_f^t(\sigma, h)$ and $U_f^t(\hat{\sigma}, h)$ are independent of σ_f and $\hat{\sigma}_f$, and since $\sigma_{\tilde{f}} = \hat{\sigma}_{\tilde{f}}$ for all $\tilde{f} \neq f$, it follows that $U_f^t(\sigma, h) = U_f^t(\hat{\sigma}, h)$. For the remainder, suppose $f \in F_a(\sigma, h)$.

Case 1: $n(\sigma, h) = 1$. Lemma A.9 implies that the sole active faction cooperates at any subgame perfect Nash equilibrium. This is precisely the behavior that σ prescribes for a sole active faction. Hence there is no profitable deviation and $U_f^t(\sigma, h) \geq U_f^t(\hat{\sigma}, h)$.

Case 2: $n(\sigma, h) \geq 2$. Under σ, this implies $m^t(\sigma, h) = 0$ and thus $y_f(h) = d$. Since $\hat{\sigma} \neq \sigma$ and $\hat{\sigma}$ agrees with σ at all histories $\tilde{h} \neq h$ and for all $\tilde{f} \neq f$, it must be that $\hat{y}_f(h) \neq y_f(h)$, and hence $\hat{y}_f(h) = c$. Note that $n(\sigma, h) = n(\hat{\sigma}, h)$ since the number of active factions at h is determined by the participation decisions made at $t = 0$, which are recorded in h and thus independent of the strategy profile. Denote $n := n(\sigma, h) = n(\hat{\sigma}, h)$ for brevity. The expected present values are therefore

$$U_f^t(\sigma, h) = -R + \Big(1 - p(n, 0)\Big) \left(\frac{1}{n}\,\Pi_1 + \frac{n-1}{n}\,\Pi_0\right) + p(n, 0)\,\beta\,U_f^{t+1}(\sigma, h)$$

and

$$U_f^t(\hat{\sigma}, h) = -R + \Big(1 - p(n, 1)\Big)\,\Pi_0 + p(n, 1)\,\beta\,U_f^{t+1}(\hat{\sigma}, h).$$

Since $n \geq 2$, we have $p(n, 0) = p(n, 1)$. Furthermore, since participation is fixed at $t = 0$, the number of active factions remains $n \geq 2$ at all $\tilde{t} > t$ and $\tilde{h} \in H^{\tilde{t}}$ succeeding h. Since both strategies agree at all $\tilde{h} \neq h$, both prescribe $m^{\tilde{t}}(\sigma, \tilde{h}) = m^{\tilde{t}}(\hat{\sigma}, \tilde{h}) = 0$ at all such histories, so the continuation values coincide: $U_f^{t+1}(\sigma, h) = U_f^{t+1}(\hat{\sigma}, h)$. Together with $\Pi_1 > \Pi_0$, this yields $U_f^t(\sigma, h) > U_f^t(\hat{\sigma}, h)$. $\square$

Lemma A.10. ***Bounds on defection equilibrium***
Let

$$\bar{R}^D = \max\{\bar{R}_1^D, \bar{R}_2^D\},$$

where

$$\bar{R}_1^D = \frac{\Pi_1}{3+\zeta_d} - \frac{2\Pi_0}{\zeta_d(1-\beta)+2}\frac{1}{3+\zeta_d},$$

$$\bar{R}_2^D = \frac{\Pi_1}{\zeta_d+2} - \frac{\Pi_0\left((1-\beta)(\zeta_d-\zeta_c)+1\right)}{\zeta_c(1-\beta)+1}\frac{1}{\zeta_d+2}.$$

Then the following holds:

A.10.1 $R \leq \bar{R}^D \Rightarrow n^D \geq 2$.

A.10.2 $R > \bar{R}^D \Rightarrow n^D \leq 2$.

A.10.3 $n^D < 3 \Rightarrow R > \bar{R}_1^D$.

Proof: Recall that

$$\Delta V^D(n) = V_d^D(n+1) - V_i^D(n),$$

and that there is a unique $\tilde{n}^D > 0$ such that $\Delta V^D(\tilde{n}^D) = 0$, with ΔV^D decreasing at $\tilde{n}^D$ (see Lemma A.4).

A.10.1: By definition of n^D, $n^D \geq 2$ if $\tilde{n}^D \geq 1$. Since ΔV^D crosses zero from above at the unique $\tilde{n}^D$, it follows that

$$\Delta V^D(1) \geq 0 \Rightarrow \tilde{n}^D \geq 1 \Rightarrow n^D \geq 2.$$

Substituting and rearranging yields

$$\Delta V^D(1) \geq 0 \quad \iff \quad R \leq \frac{\Pi_1}{\zeta_d+2} - \frac{\Pi_0}{\zeta_d(1-\beta)+1}\frac{1}{\zeta_d+2}.$$

We show that this bound strictly exceeds $\bar{R}^D$. First,

$$\frac{\Pi_1}{\zeta_d+2} - \frac{\Pi_0}{\zeta_d(1-\beta)+1}\frac{1}{\zeta_d+2} > \bar{R}_2^D \quad \iff \quad (1-\beta)(\zeta_d-\zeta_c) > 0,$$

which is always true. Second, we compare the bound with $\bar{R}_1^D$. Subtracting

$\bar{R}_1^D$ from the bound and collecting terms over the common factor $\frac{1}{\zeta_d+2}$, the comparison reduces to the sign of

$$\Pi_1 + \Pi_0 \frac{\zeta_d(1-\beta)(\zeta_d+1)-2}{\Big(\zeta_d(1-\beta)+1\Big)\Big(\zeta_d(1-\beta)+2\Big)}.$$

To sign this expression, consider the fraction

$$\frac{\zeta_d(1-\beta)(\zeta_d+1)-2}{\Big(\zeta_d(1-\beta)+1\Big)\Big(\zeta_d(1-\beta)+2\Big)}.$$

The denominator is strictly positive, since $\beta \in (0,1)$ and $\zeta_d > 0$, so the sign of the fraction is determined by the numerator $\zeta_d(1-\beta)(\zeta_d+1)-2$. Since $\zeta_d(1-\beta)(\zeta_d+1) \geq 0$, the numerator is bounded below by -2. The denominator satisfies

$$\Big(\zeta_d(1-\beta)+1\Big)\Big(\zeta_d(1-\beta)+2\Big) \geq 2,$$

with equality as $\zeta_d(1-\beta) \to 0$, so the fraction is bounded below by -1. Together with $\Pi_1 > \Pi_0$, this gives

$$\Pi_1 + \Pi_0 \frac{\zeta_d(1-\beta)(\zeta_d+1)-2}{\Big(\zeta_d(1-\beta)+1\Big)\Big(\zeta_d(1-\beta)+2\Big)} \geq \Pi_1 - \Pi_0 > 0,$$

and thus

$$\frac{\Pi_1}{\zeta_d+2} - \frac{\Pi_0}{\zeta_d(1-\beta)+1}\frac{1}{\zeta_d+2} > \bar{R}_1^D.$$

Combining the two comparisons,

$$\frac{\Pi_1}{\zeta_d+2} - \frac{\Pi_0}{\zeta_d(1-\beta)+1}\frac{1}{\zeta_d+2} > \max\{\bar{R}_1^D, \bar{R}_2^D\} = \bar{R}^D.$$

Thus,

$$R \leq \bar{R}^D \Rightarrow R \leq \frac{\Pi_1}{\zeta_d+2} - \frac{\Pi_0}{\zeta_d(1-\beta)+1}\frac{1}{\zeta_d+2} \Rightarrow \Delta V^D(1) \geq 0 \Rightarrow \tilde{n}^D \geq 1 \Rightarrow n^D \geq 2.$$

A.10.2: By definition of n^D, $n^D \leq 2$ if $\tilde{n}^D < 2$. Since ΔV^D crosses zero from above at the unique $\tilde{n}^D$, it follows that

$$\Delta V^D(2) < 0 \Rightarrow \tilde{n}^D < 2 \Rightarrow n^D \leq 2.$$

Substituting and rearranging yields

$$\Delta V^D(2) < 0 \quad \Longleftrightarrow \quad R > \frac{\Pi_1}{3+\zeta_d} - \frac{2\Pi_0}{\zeta_d(1-\beta)+2}\,\frac{1}{3+\zeta_d} = \bar{R}_1^D.$$

Hence,

$$R > \bar{R}^D \Rightarrow R > \bar{R}_1^D \Rightarrow \Delta V^D(2) < 0 \Rightarrow \tilde{n}^D < 2 \Rightarrow n^D \leq 2.$$

A.10.3: By definition of n^D, $n^D < 3$ if and only if $\tilde{n}^D < 2$. By the reasoning for Part A.10.2, $\tilde{n}^D < 2$ if and only if $R > \bar{R}_1^D$. Hence, $n^D < 3 \Rightarrow R > \bar{R}_1^D$. $\square$

Proof of Proposition 1

To prove the proposition, we define the following:

$$\bar{R}^D = \max\{\bar{R}_1^D, \bar{R}_2^D\},$$

where

$$\bar{R}_1^D = \frac{\Pi_1}{3+\zeta_d} - \frac{2\Pi_0}{\zeta_d(1-\beta)+2}\frac{1}{3+\zeta_d}$$

and

$$\bar{R}_2^D = \frac{\Pi_1}{\zeta_d+2} - \frac{\Pi_0\left((1-\beta)(\zeta_d-\zeta_c)+1\right)}{\zeta_c(1-\beta)+1}\frac{1}{\zeta_d+2}.$$

Note that $\Pi_1 > \Pi_0$ and $\frac{2}{\zeta_d\,(1-\beta)+2} < 1$ imply $\bar{R}_1^D > 0$ and thus $\bar{R}^D > 0$.

If: We first show that if $R \leq \bar{R}^D$, then there is an opposition equilibrium as described. We do so by providing an explicit example. Suppose $R \leq \bar{R}^D$. Note that $R \leq \bar{R}^D \Rightarrow R < \frac{\Pi_1}{\zeta_d+1} \Rightarrow \tilde{n}^D \geq 0$ exists. Let $F^D \subseteq F$ with $|F^D| = n^D$ and define $\sigma = (\sigma_f)_{f\in F}$ with $\sigma_f = (x_f, y_f)$ as follows:

1. For all $f \in F$:

$$x_f = a \Leftrightarrow f \in F^D$$

2. For all $t \geq 1$, $h \in H^t$, and $f \in F_a(\sigma, h)$:

$$y_f(h) = d \Leftrightarrow n(\sigma, h) \geq 2$$

By construction, $n(\sigma, \emptyset) = n^D$, which follows from Condition 1 and $|F^D| = n^D$. Since Lemma A.10 and $R \leq \bar{R}^D$ imply $n^D \geq 2$, we have $n(\sigma, h) \geq 2$ at every history on the path of play. Condition 2 then implies $y_f(h) = d$ for all active f and all on-path histories h, and hence $m^t(\sigma, \emptyset) = 0$ for all $t \geq 1$.

We now show that σ is an opposition equilibrium. Since x_f is constant and $y_f(h)$ depends only on $n(\sigma, h)$, which is itself determined by $\{x_{\hat{f}}\}_{\hat{f}\in F}$, Condition 2 of Definition 1 holds. It remains to show that σ is a subgame perfect Nash equilibrium, for which we employ the one-shot deviation principle.

First, we investigate whether any faction has an incentive to deviate on the interaction dimension. Consider any $t \geq 1$, $h \in H^t$, and $f \in F_a(\sigma, h)$. By construction of σ:

$$n(\sigma, h) \geq 2 \Leftrightarrow y_{\tilde{f}}(h) = d \;\forall \tilde{f} \in F_a(\sigma, h) \Leftrightarrow m^t(\sigma, h) = 0.$$

Lemma 3 then implies that no faction f has a profitable one-shot deviation on the interaction dimension at (t, h). Since t and h were arbitrary, this holds for all $t \geq 1$ and $h \in H^t$.

Second, we show that no faction has an incentive to deviate on the participation dimension. Consider any $f \in F$ and let $\hat{\sigma}$ be such that $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ for all $\tilde{f} \neq f$, with $\hat{\sigma}_f = (\hat{x}_f, y_f)$, that is, $\hat{\sigma}$ differs from σ only in the participation decision of f.

Case 1: $n^D \geq 3$, or $n^D \geq 2$ and $x_f = i$. Since only f changes its participation decision, $n(\hat{\sigma}, \emptyset) \in \{n^D - 1, n^D + 1\}$. In the first case, $n(\hat{\sigma}, \emptyset) \geq n^D - 1 \geq 2$. In the second case, the deviation is from i to a, so $n(\hat{\sigma}, \emptyset) = n^D + 1 \geq 3 \geq 2$. In both cases, by construction of σ and since $\hat{\sigma}$ and σ coincide on $y_{\tilde{f}}$ for all $\tilde{f}$, it follows that $y_{\tilde{f}}(h) = d$ for all $\tilde{f} \in F_a(\hat{\sigma}, h)$ and all histories h along the path of play. Consequently, $m^t(\sigma, \emptyset) = m^t(\hat{\sigma}, \emptyset) = 0$ for all $t \geq 1$. Lemma 1 then implies $U^0_f(\hat{\sigma}, \emptyset) \leq U^0_f(\sigma, \emptyset)$.

Case 2: $n^D = 2$ and $x_f = a$. Here $f \in F^D$ and the deviation is from a to i, so $n(\hat{\sigma}, \emptyset) = n^D - 1 = 1$. Since $n(\hat{\sigma}, h) = 1 < 2$ along the path of play, Condition 2 implies $y_{\tilde{f}}(h) = c$ for all active $\tilde{f}$ under $\hat{\sigma}$. Thus, by becoming inactive, f leaves a single faction behind that cooperates every period. Consequently, $U^0_f(\hat{\sigma}, \emptyset) = V^C_i(1)$ while $U^0_f(\sigma, \emptyset) = V^D_d(2)$. Substituting and rearranging yields that the condition $U^0_f(\sigma, \emptyset) \geq U^0_f(\hat{\sigma}, \emptyset)$ reduces to $R \leq \bar{R}^D_2$. Since $n^D < 3$ implies $R > \bar{R}^D_1$ (see Lemma A.10), and $R \leq \bar{R}^D$ holds by assumption, it follows that $R \leq \bar{R}^D_2$, and hence $U^0_f(\sigma, \emptyset) \geq U^0_f(\hat{\sigma}, \emptyset)$.

Since f was chosen arbitrarily and no profitable one-shot deviation exists on either the participation or interaction dimension, σ is a subgame perfect Nash equilibrium and hence an opposition equilibrium.

Only if: Next, we show that no opposition equilibrium as described exists if $R > \bar{R}^D$. We prove this by contradiction: suppose σ is an opposition equilibrium as described and $R > \bar{R}^D$. $R > \bar{R}^D$ implies $R > \bar{R}^D_1$. In combination with Lemma A.10, it follows that $n^D \leq 2$. By definition, $\tilde{n}^D \geq 0$ and, thus, $n^D \in \{1, 2\}$.

Case 1: $n^D = 1$. Hence, $n(\sigma, h) = n^D$ for all histories h along the path of play. Lemma A.9 then implies $m^t(\sigma, h) = 1$ for all h along the path of play, and hence $m^t(\sigma, \emptyset) = 1$ for all $t \geq 1$. Thus $n^D = 1$ is not possible, as it yields a contradiction.

Case 2: $n^D = 2$. Consider any $f \in F_a(\sigma, \emptyset)$. Under σ, f is active and $n^D = 2$.

Moreover, along the path of play, all active factions defect since $m^t(\sigma, \emptyset) = 0$ for all $t \geq 1$. Hence, $U_f^0(\sigma, \emptyset) = V_d^D(2)$. Now consider any $\hat{\sigma}$ that differs from σ only in the participation decision of f, so that f is inactive under $\hat{\sigma}$ and $n(\hat{\sigma}, \emptyset) = 1$. Since σ and $\hat{\sigma}$ coincide on all interaction decisions and σ is a subgame perfect Nash equilibrium, Lemma A.9 implies $n(\hat{\sigma}, h) = 1 \Rightarrow m^t(\hat{\sigma}, h) = 1$ for all $t \geq 1$ and $h \in H^t$. Since $n(\hat{\sigma}, h) = 1$ for all h along the path of play under $\hat{\sigma}$, it follows that $m^t(\hat{\sigma}, h) = 1$ for all such h. Thus, $U_f^0(\hat{\sigma}, \emptyset) = V_i^C(1)$. Since σ is a subgame perfect Nash equilibrium, it must hold that $U_f^0(\hat{\sigma}, \emptyset) \leq U_f^0(\sigma, \emptyset)$. Substituting $V_i^C(1)$ and $V_d^D(2)$ and rearranging yields that this condition is equivalent to $R \leq \bar{R}_2^D$, which contradicts $R > \bar{R}^D \geq \bar{R}_2^D$. Hence, there is no opposition equilibrium as described whenever $R > \bar{R}^D$. $\square$

A.2 Formal analysis of the cooperation equilibrium

Lemma A.11. ***Comparative statics of*** $V_i^C(n)$
The inactivity value $V_i^C(n)$ *satisfies the following comparative statics:*

A.11.1 $V_i^C(n)$ *is strictly increasing and strictly concave in* n.

A.11.2 $V_i^C(n)$ *is strictly increasing in* Π_0.

A.11.3 $V_i^C(n)$ *is strictly increasing in* β.

A.11.4 $V_i^C(n)$ *is strictly decreasing in* ζ_c.

Proof: The closed form

$$V_i^C(n) = \frac{n\Pi_0}{\zeta_c(1-\beta)+n}$$

coincides with the closed form of $V_i^D(n)$ after replacing ζ_d with ζ_c. Since $\zeta_c > 0$, the denominator remains strictly positive, and every step of the proof of Lemma A.1 carries over under this substitution. In particular, the derivatives with respect to n, Π_0, β, and ζ_c take the same form with ζ_c in place of ζ_d, and their signs are unchanged. $\square$

Lemma A.12. ***Comparative statics of*** $V_c^C(n)$
The activity value $V_c^C(n)$ *satisfies the following comparative statics:*

A.12.1 $V_c^C(n)$ *is strictly increasing in* Π_1*.*

A.12.2 $V_c^C(n)$ *is increasing in* Π_0*, and strictly increasing in* Π_0 *whenever* $n > 1$*.*

A.12.3 $V_c^C(n)$ *is strictly decreasing in* R*.*

A.12.4 If $V_c^C(n) > 0$*, then* $V_c^C(n)$ *is strictly increasing in* β*.*

A.12.5 $V_c^C(n)$ *is strictly decreasing in* ζ_c*.*

Proof: The closed form

$$V_c^C(n) = \frac{\Delta\Pi + n\Pi_0 - R(\zeta_c + n)}{\zeta_c(1-\beta) + n}$$

coincides with the closed form of $V_d^D(n)$ after replacing ζ_d with ζ_c. Since $\zeta_c > 0$, the denominator remains strictly positive, and every step of the proof of Lemma A.2 carries over under this substitution. In particular, the derivatives with respect to Π_1, Π_0, R, β, and ζ_c take the same form with ζ_c in place of ζ_d, and their signs are unchanged, with the condition $V_d^D(n) > 0$ in the β part replaced by $V_c^C(n) > 0$. □

Lemma A.13. ***Monotonicity and curvature of*** $V_c^C(n)$ ***in*** n

A.13.1 If
$$\zeta_c\Big((1-\beta)\Pi_0 + \beta R\Big) > \Delta\Pi,$$
then $V_c^C(n)$ *is strictly increasing and strictly concave in* n*.*

A.13.2 If
$$\zeta_c\Big((1-\beta)\Pi_0 + \beta R\Big) < \Delta\Pi,$$
then $V_c^C(n)$ *is strictly decreasing and strictly convex in* n*.*

Proof: The closed form

$$V_c^C(n) = \frac{\Delta\Pi - R\zeta_c + n(\Pi_0 - R)}{\zeta_c(1-\beta) + n}$$

coincides with the closed form of $V_d^D(n)$ after replacing ζ_d with ζ_c. Since $\zeta_c > 0$, the denominator remains strictly positive, and every step of the proof of Lemma A.3 carries over under this substitution. In particular, the first and second derivatives with respect to n take the same form with ζ_c in place of ζ_d, so their signs are governed by the sign of $\zeta_c\Big((1-\beta)\Pi_0 + \beta R\Big) - \Delta\Pi$, exactly as stated in the two parts. $\square$

Lemma A.14. ***Properties of*** $\tilde{n}_a^C$

A.14.1 The threshold $\tilde{n}_a^C \geq 0$ *is well defined if and only if*

$$R \leq \frac{\Pi_1}{\zeta_c + 1}.$$

A.14.2 If $\tilde{n}_a^C \geq 0$ *exists, it is unique.*

A.14.3 If $\tilde{n}_a^C \geq 0$ *exists, then*

$$\Delta V_a^{C\prime}(\tilde{n}_a^C) < 0.$$

A.14.4

$$\tilde{n}_a^C = \frac{\Delta\Pi - R(\zeta_c(2-\beta)+1) + \sqrt{\begin{array}{r}\Big(\Delta\Pi - R(\zeta_c(2-\beta)+1)\Big)^2 \\ + 4R\zeta_c(1-\beta)\Big(\Pi_1 - R(\zeta_c+1)\Big)\end{array}}}{2R}.$$

Proof: Recall that

$$\Delta V_a^C(n) = V_c^C(n+1) - V_i^C(n),$$

where $V_c^C(n+1)$ and $V_i^C(n)$ coincide with $V_d^D(n+1)$ and $V_i^D(n)$ after replacing ζ_d with ζ_c. The threshold $\tilde{n}_a^C$ solves $\Delta V_a^C(\tilde{n}_a^C) = 0$, and clearing denominators yields the quadratic of the proof of Lemma A.4 with ζ_c in place of ζ_d,

$$-Rn^2 + \Big(\Delta\Pi - R(\zeta_c(2-\beta)+1)\Big)n + \zeta_c(1-\beta)\Big(\Pi_1 - R(\zeta_c+1)\Big) = 0.$$

Since $\zeta_c > 0$ and $\beta \in (0,1)$, every step of that proof carries over under this substitution. The sign of the constant term is governed by $\Pi_1 - R(\zeta_c+1)$, so the existence bound becomes $R \leq \Pi_1/(\zeta_c+1)$. In the converse direction, the chain $R(\zeta_c(2-\beta)+1) > R(\zeta_c+1) > \Pi_1 > \Delta\Pi$ holds whenever $R > \Pi_1/(\zeta_c+1)$, so the root-sum argument is unchanged. The uniqueness argument via the product of roots and the slope argument via evaluation at the larger root are likewise unchanged, with ζ_c in place of ζ_d throughout. Finally, the closed-form solution also mirrors the earlier one. $\square$

Proof of Lemma 4

Consider any $\sigma \in C_x$ such that $n(\sigma, \emptyset) = n_a^C$. Moreover, consider any $\hat{\sigma} \in C_x$ such that $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ for all $\tilde{f} \neq f$. Let $\hat{\sigma}_{\tilde{f}} = (\hat{x}_{\tilde{f}}, \hat{y}_{\tilde{f}})$ and $\sigma_{\tilde{f}} = (x_{\tilde{f}}, y_{\tilde{f}})$ for any $\tilde{f}$.

Case 1: $\hat{x}_f = x_f$. Then no faction changes its participation decision under $\hat{\sigma}$ as compared to σ, that is, $\{x_{\tilde{f}}\}_{\tilde{f}\in F} = \{\hat{x}_{\tilde{f}}\}_{\tilde{f}\in F}$, and thus $n(\sigma, \emptyset) = n(\hat{\sigma}, \emptyset)$. Since $\hat{\sigma}, \sigma \in C_x$, $m^t(\sigma, \emptyset) = m^t(\hat{\sigma}, \emptyset) = n(\sigma, \emptyset)$ for all $t \geq 1$. All active factions always cooperate along the path of play under $\hat{\sigma}$ and σ. Since no faction changes its on-path play at $\hat{\sigma}$ as compared to σ, the expected present values of f coincide: $U_f^0(\sigma, \emptyset) = U_f^0(\hat{\sigma}, \emptyset)$.

Case 2: $x_f = a$ and $\hat{x}_f = i$. Faction f being inactive under $\hat{\sigma}$ and active under σ implies $n(\hat{\sigma}, \emptyset) = n_a^C - 1$. Moreover, since $\hat{\sigma}, \sigma \in C_x$, $m^t(\sigma, \emptyset) = n_a^C$ and $m^t(\hat{\sigma}, \emptyset) = n_a^C - 1$ for all $t \geq 1$. Thus, the resulting expected present values are $U_f^0(\sigma, \emptyset) = V_c^C(n_a^C)$ and $U_f^0(\hat{\sigma}, \emptyset) = V_i^C(n_a^C - 1)$. By definition of $\tilde{n}_a^C$, $V_c^C(\tilde{n}_a^C + 1) - V_i^C(\tilde{n}_a^C) = 0$. By Lemma A.14, $V_c^C(x) - V_i^C(x-1)$ is decreasing in x, and by definition of n_a^C, $n_a^C \leq \tilde{n}_a^C + 1$. Hence, $V_c^C(n_a^C) - V_i^C(n_a^C - 1) \geq 0$ and thus $U_f^0(\sigma, \emptyset) \geq U_f^0(\hat{\sigma}, \emptyset)$.

Case 3: $x_f = i$ and $\hat{x}_f = a$. Faction f being active under $\hat{\sigma}$ and inactive under σ implies $n(\hat{\sigma}, \emptyset) = n_a^C + 1$. Moreover, since $\hat{\sigma}, \sigma \in C_x$, $m^t(\sigma, \emptyset) = n_a^C$ and $m^t(\hat{\sigma}, \emptyset) = n_a^C + 1$ for all $t \geq 1$. Thus, the resulting expected present values are $U_f^0(\sigma, \emptyset) = V_i^C(n_a^C)$ and $U_f^0(\hat{\sigma}, \emptyset) = V_c^C(n_a^C + 1)$. By definition of $\tilde{n}_a^C$, $V_c^C(\tilde{n}_a^C + 1) - V_i^C(\tilde{n}_a^C) = 0$. By Lemma A.14, $V_c^C(x+1) - V_i^C(x)$ is decreasing in x, and by definition of n_a^C, $n_a^C \geq \tilde{n}_a^C$. Hence, $V_c^C(n_a^C + 1) - V_i^C(n_a^C) \leq 0$ and thus $U_f^0(\sigma, \emptyset) \geq U_f^0(\hat{\sigma}, \emptyset)$. $\square$

Lemma A.15. ***Comparative statics of*** $\tilde{n}_a^C$
Suppose that $\tilde{n}_a^C > 0$ *exists. Then the cooperation-entry threshold satisfies the following comparative statics:*

A.15.1 $\tilde{n}_a^C$ *is strictly increasing in* Π_1.

A.15.2 $\tilde{n}_a^C$ *is strictly decreasing in* Π_0.

A.15.3 $\tilde{n}_a^C$ *is strictly decreasing in* R.

A.15.4 $\tilde{n}_a^C$ *is strictly decreasing in* β.

Proof: Recall that

$$\Delta V_a^C(n) = V_c^C(n+1) - V_i^C(n),$$

where $V_c^C(n+1)$ and $V_i^C(n)$ coincide with $V_d^D(n+1)$ and $V_i^D(n)$ after replacing ζ_d with ζ_c. The cooperation-entry threshold $\tilde{n}_a^C$ solves $\Delta V_a^C(\tilde{n}_a^C) = 0$, and by Lemma A.14, $\partial \Delta V_a^C / \partial \tilde{n}_a^C < 0$ at the equilibrium intersection. Since $\zeta_c > 0$, every step of the proof of Lemma A.5 carries over under this substitution: the implicit function theorem gives that the sign of $\partial \tilde{n}_a^C / \partial \theta$ equals the sign of $\partial \Delta V_a^C / \partial \theta$ for any scalar parameter θ, and the derivatives of ΔV_a^C with respect to Π_1, Π_0, R, and β take the same form as in that proof with ζ_c in place of ζ_d, so their signs are unchanged. $\square$

Lemma A.16. ***Comparative statics of*** $\tilde{n}_a^C$ ***in*** ζ_c

Suppose that $\tilde{n}_a^C > 0$ *exists. Then*

$$\frac{\partial \tilde{n}_a^C}{\partial \zeta_c} \leq 0 \quad \iff \quad \zeta_c \geq \kappa,$$

where $\kappa \in [0, \infty)$ *is defined by*

$$\kappa := \begin{cases} 0 & \text{if } \Pi_1 > R + (2-\beta)\Pi_0, \\ \dfrac{\Pi_1 - R - \dfrac{2-\beta}{\beta}\Pi_0 \left(1 - \sqrt{\dfrac{\Pi_0 + \beta R - \beta \Pi_1}{\Pi_0}}\right)}{\beta R} & \text{otherwise.} \end{cases}$$

Proof: Recall the closed-form solution of $\tilde{n}_a^C$ in Lemma A.14. Every step of the proof of Lemma A.6 carries over when substituting ζ_d with ζ_c: the derivative of the larger root with respect to ζ_c takes the same form, the case split is governed by the sign of $R\beta^2\zeta_c + R\beta + (2-\beta)\Pi_0 - \beta\Pi_1$, and the squaring argument yields the same condition $\zeta_c \geq \kappa$. Note that the cutoff κ does not depend on the structural advantage, so it is identical to the cutoff in Lemma A.6. □

Lemma A.17. ***Comparative statics of*** n_a^C

A.17.1 n_a^C *is increasing in* Π_1.

A.17.2 n_a^C *is decreasing in* Π_0.

A.17.3 n_a^C *is decreasing in* R.

A.17.4 n_a^C *is decreasing in* β.

A.17.5 n_a^C *is decreasing in* ζ_c *if* $\zeta_c \geq \kappa$ *and increasing in* ζ_c *if* $\zeta_c \leq \kappa$*, where* κ *is defined as in Lemma A.6.*

Proof: Recall that n_a^C is the smallest integer larger than $\tilde{n}_a^C$. By analogous reasoning as for n^D in Lemma A.7, Lemma A.17 then follows directly from Lemma A.15 and Lemma A.16. □

Lemma A.18. ***Repression and*** $\tilde{n}_a^C$

Suppose that $\tilde{n}_a^C \geq 0$ *exists. Then* $\tilde{n}_a^C$ *satisfies the following properties:*

A.18.1 $\lim_{R \to 0^+} \tilde{n}_a^C = \infty$.

A.18.2 $\lim_{R \to \Pi_1/(\zeta_c+1)^-} \tilde{n}_a^C = 0$.

Proof: The proof mirrors that of Lemma A.8, replacing $\tilde{n}^D$ with $\tilde{n}_a^C$, the two functions differing only with respect to ζ_d and ζ_c. $\square$

Proof of Lemma 5

The proof is analogous to that of Lemma 2, drawing inferences on n_a^C through $\tilde{n}_a^C$ (using Lemma A.18) and Lemma A.17, in place of n^D through $\tilde{n}^D$ (using Lemma A.8) and Lemma A.7. $\square$

Lemma A.19. ***Comparative statics of*** $V_d^C(n)$

The defection value $V_d^C(n)$ *satisfies the following comparative statics:*

A.19.1 $V_d^C(n)$ *is strictly increasing and strictly concave in* n*.*

A.19.2 $V_d^C(n)$ *is strictly increasing in* Π_1*.*

A.19.3 $V_d^C(n)$ *is strictly decreasing in* R*.*

A.19.4 If $V_d^C(n) > 0$*, then* $V_d^C(n)$ *is strictly increasing in* β*.*

A.19.5 $V_d^C(n)$ *is strictly decreasing in* ζ_d*.*

Proof: Recall the closed form

$$V_d^C(n) = \frac{n\Pi_1 - R(\zeta_d + n)}{\zeta_d(1-\beta) + n}.$$

Throughout, the denominator $\zeta_d(1-\beta) + n$ is strictly positive, since $\zeta_d > 0$, $\beta \in (0,1)$, and $n \geq 1$.

A.19.1: Differentiating with respect to n gives

$$\frac{\partial V_d^C(n)}{\partial n} = \frac{\zeta_d\Big((1-\beta)\Pi_1 + \beta R\Big)}{\Big(\zeta_d(1-\beta) + n\Big)^2} > 0,$$

and

$$\frac{\partial^2 V_d^C(n)}{\partial n^2} = -2\frac{\zeta_d\Big((1-\beta)\Pi_1 + \beta R\Big)}{\Big(\zeta_d(1-\beta) + n\Big)^3} < 0.$$

Hence $V_d^C(n)$ is strictly increasing and strictly concave in n.

A.19.2: Differentiating with respect to Π_1 gives

$$\frac{\partial V_d^C(n)}{\partial \Pi_1} = \frac{n}{\zeta_d(1-\beta) + n} > 0.$$

Hence $V_d^C(n)$ is strictly increasing in Π_1.

A.19.3: Differentiating with respect to R gives

$$\frac{\partial V_d^C(n)}{\partial R} = -\frac{\zeta_d + n}{\zeta_d(1-\beta) + n} < 0.$$

Hence $V_d^C(n)$ is strictly decreasing in R.

A.19.4: Differentiating with respect to β gives

$$\frac{\partial V_d^C(n)}{\partial \beta} = \frac{\zeta_d\Big(n\Pi_1 - R(\zeta_d + n)\Big)}{\Big(\zeta_d(1-\beta) + n\Big)^2} = \frac{\zeta_d V_d^C(n)}{\zeta_d(1-\beta) + n}.$$

This expression is strictly positive whenever $V_d^C(n) > 0$. Hence $V_d^C(n)$ is strictly increasing in β whenever $V_d^C(n) > 0$.

A.19.5: Differentiating with respect to ζ_d gives

$$\frac{\partial V_d^C(n)}{\partial \zeta_d} = -\frac{n\Big((1-\beta)\Pi_1 + \beta R\Big)}{\Big(\zeta_d(1-\beta) + n\Big)^2} < 0.$$

Hence $V_d^C(n)$ is strictly decreasing in ζ_d. □

Lemma A.20. ***Properties of*** $\tilde{n}_c^C$

The cooperation-interaction threshold $\tilde{n}_c^C$ *satisfies the following properties:*

A.20.1 The threshold $\tilde{n}_c^C$ *is always well-defined, strictly positive, and unique.*

A.20.2 $\Delta V_c^{C\prime}(\tilde{n}_c^C) < 0$.

A.20.3

$$\tilde{n}_c^C = \frac{\begin{array}{c}\Delta\Pi + \beta R(\zeta_d - \zeta_c) + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1) \\ + \sqrt{\begin{array}{l}\Big(\Delta\Pi + \beta R(\zeta_d - \zeta_c) + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1)\Big)^2 \\ \quad + 4(1-\beta)\zeta_d \Delta\Pi^2\end{array}}\end{array}}{2\Delta\Pi}.$$

Proof: Recall that

$$\Delta V_c^C(n) = V_c^C(n) - V_d^C(n),$$

where

$$V_c^C(n) = \frac{\Delta\Pi + n\Pi_0 - R(\zeta_c + n)}{\zeta_c(1-\beta) + n}, \qquad V_d^C(n) = \frac{n\Pi_1 - R(\zeta_d + n)}{\zeta_d(1-\beta) + n}.$$

The threshold $\tilde{n}_c^C$ solves $\Delta V_c^C(\tilde{n}_c^C) = 0$. Clearing denominators, this is equivalent to $Q(\tilde{n}_c^C) = 0$, where

$$\begin{aligned} Q(n) := & -\Delta\Pi\, n^2 + \Big(\Pi_0(\zeta_d(1-\beta) - 1) - \Pi_1(\zeta_c(1-\beta) - 1) + \beta R(\zeta_d - \zeta_c)\Big)n \\ & + \zeta_d(1-\beta)\Delta\Pi. \end{aligned} \tag{27}$$

A.20.1: The constant term $\zeta_d(1-\beta)\Delta\Pi$ is strictly positive. Since the coefficient on n^2 is $-\Delta\Pi < 0$, the parabola Q opens downward and is positive at $n = 0$, and since $Q(n) \to -\infty$ as $n \to +\infty$, it must cross zero for some $n > 0$. Hence $\tilde{n}_c^C > 0$ exists. By similar reasoning, the other root is strictly negative, so $\tilde{n}_c^C$ is the unique non-negative root.

A.20.2: We rewrite

$$\Delta V_c^C(n) = \frac{Q(n)}{(\zeta_c(1-\beta) + n)(\zeta_d(1-\beta) + n)},$$

with $Q(n)$ as defined in (27). Since $Q(\tilde{n}_c^C) = 0$ and the denominator is strictly positive at $\tilde{n}_c^C$, it follows that

$$\Delta V_c^{C\prime}(\tilde{n}_c^C) = \frac{Q'(\tilde{n}_c^C)}{(\zeta_c(1-\beta)+\tilde{n}_c^C)(\zeta_d(1-\beta)+\tilde{n}_c^C)}.$$

Since $Q(n)$ is a downward-opening parabola and $\tilde{n}_c^C$ is the larger of its two roots, $Q'(\tilde{n}_c^C) < 0$ and thus $\Delta V_c^{C\prime}(\tilde{n}_c^C) < 0$.

A.20.3: Applying the quadratic formula to (27), the two roots are

$$\begin{aligned} n &= \frac{\begin{array}{l} -\Big(\Pi_0(\zeta_d(1-\beta)-1) - \Pi_1(\zeta_c(1-\beta)-1) + \beta R(\zeta_d - \zeta_c)\Big) \\ \quad \pm \sqrt{\begin{array}{l}\Big(\Pi_0(\zeta_d(1-\beta)-1) - \Pi_1(\zeta_c(1-\beta)-1) + \beta R(\zeta_d - \zeta_c)\Big)^2 \\ \quad + 4(1-\beta)\zeta_d \Delta\Pi^2\end{array}} \end{array}}{-2\Delta\Pi} \\ &= \frac{\begin{array}{l} \Pi_0(\zeta_d(1-\beta)-1) - \Pi_1(\zeta_c(1-\beta)-1) + \beta R(\zeta_d - \zeta_c) \\ \quad \mp \sqrt{\begin{array}{l}\Big(\Pi_0(\zeta_d(1-\beta)-1) - \Pi_1(\zeta_c(1-\beta)-1) + \beta R(\zeta_d - \zeta_c)\Big)^2 \\ \quad + 4(1-\beta)\zeta_d \Delta\Pi^2\end{array}} \end{array}}{2\Delta\Pi}. \end{aligned}$$

The unique non-negative cooperation-interaction threshold must be the larger one and thus the root taking the positive square root,

$$\tilde{n}_c^C = \frac{\begin{array}{l} \Delta\Pi + \beta R(\zeta_d - \zeta_c) + (1-\beta)(\zeta_d\Pi_0 - \zeta_c\Pi_1) \\ \quad + \sqrt{\begin{array}{l}\Big(\Delta\Pi + \beta R(\zeta_d - \zeta_c) + (1-\beta)(\zeta_d\Pi_0 - \zeta_c\Pi_1)\Big)^2 \\ \quad + 4(1-\beta)\zeta_d \Delta\Pi^2\end{array}} \end{array}}{2\Delta\Pi}.$$

□

Proof of Lemma 6

Consider any σ as described, and let $\sigma_f = (x_f, y_f)$ for all $f \in F$. Note that the two conditions on σ are equivalent to

1. $n(\sigma, h) \leq n_c^C \Rightarrow y_f(h) = c$ and
2. $n(\sigma, h) > n_c^C \Rightarrow y_f(h) = d$

for all $f \in F_a(\sigma, h)$ and all $t \geq 1$, $h \in H^t$. In the following, consider any $f \in F$, $t \geq 1$, $h \in H^t$, and $\hat{\sigma} \neq \sigma$ such that

1. $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ for all $\tilde{f} \neq f$ and
2. $\hat{\sigma}_f(\tilde{h}) = \sigma_f(\tilde{h})$ for all $\tilde{h} \neq h$.

Let $\hat{\sigma}_{\tilde{f}} = (\hat{x}_{\tilde{f}}, \hat{y}_{\tilde{f}})$ for all $\tilde{f} \in F$. If $f \notin F_a(\sigma, h)$, then since participation is fixed at $t = 0$, f takes no actions at h or any future history. Thus $U_f^t(\sigma, h)$ and $U_f^t(\hat{\sigma}, h)$ are independent of σ_f and $\hat{\sigma}_f$, and since $\sigma_{\tilde{f}} = \hat{\sigma}_{\tilde{f}}$ for all $\tilde{f} \neq f$, it follows that $U_f^t(\sigma, h) = U_f^t(\hat{\sigma}, h)$. For the remainder, suppose $f \in F_a(\sigma, h)$. Moreover, note that $n(\sigma, h) = n(\hat{\sigma}, h)$ since the number of active factions at h is determined by the participation decisions made at $t = 0$, which are recorded in h and thus independent of the strategy profile. Denote $n := n(\sigma, h) = n(\hat{\sigma}, h)$ for brevity.

Case 1: $n \leq n_c^C$. Under σ, this implies $m^t(\sigma, h) = n$ and thus $y_f(h) = c$. Since $\hat{\sigma} \neq \sigma$ and $\hat{\sigma}$ agrees with σ at all histories $\tilde{h} \neq h$ and for all $\tilde{f} \neq f$, it must be that $\hat{y}_f(h) \neq y_f(h)$, and hence $\hat{y}_f(h) = d$. Consequently, $m^t(\hat{\sigma}, h) = n - 1$. The expected present values are therefore

$$U_f^t(\sigma, h) = -R + (1 - p(n, n)) \left(\frac{1}{n} \Pi_1 + \frac{n-1}{n} \Pi_0 \right) + p(n, n)\, \beta\, U_f^{t+1}(\sigma, h)$$

and

$$U_f^t(\hat{\sigma}, h) = -R + (1 - p(n, n-1))\, \Pi_1 + p(n, n-1)\, \beta\, U_f^{t+1}(\hat{\sigma}, h).$$

Since participation is fixed at $t = 0$, the number of active factions remains $n \leq n_c^C$ at all $\tilde{t} > t$ and $\tilde{h} \in H^{\tilde{t}}$ succeeding h. Since both strategies agree at all $\tilde{h} \neq h$, both prescribe $m^{\tilde{t}}(\sigma, \tilde{h}) = m^{\tilde{t}}(\hat{\sigma}, \tilde{h}) = n$ at all such histories, so the continuation values coincide: $U_f^{t+1}(\sigma, h) = U_f^{t+1}(\hat{\sigma}, h)$. Since under σ all active factions always cooperate at all successor histories, $U_f^{t+1}(\sigma, h) = V_c^C(n)$ and thus $U_f^{t+1}(\hat{\sigma}, h) = V_c^C(n)$. Substituting yields

$$U_f^t(\sigma, h) = V_c^C(n) = -R + (1 - p(n,n))\left(\frac{1}{n}\Pi_1 + \frac{n-1}{n}\Pi_0\right) + p(n,n)\,\beta\,V_c^C(n)$$

and

$$U_f^t(\hat{\sigma}, h) = -R + (1 - p(n, n-1))\,\Pi_1 + p(n, n-1)\,\beta\,V_c^C(n).$$

Since $p(n, n-1) = p(n, 0)$, rearranging yields

$$U_f^t(\hat{\sigma}, h) \leq U_f^t(\sigma, h) \Leftrightarrow V_c^C(n) \geq \frac{-R + (1 - p(n,0))\Pi_1}{1 - p(n,0)\beta} = V_d^C(n),$$

which holds for all $n \leq n_c^C$ by Lemma A.20. Hence $U_f^t(\hat{\sigma}, h) \leq U_f^t(\sigma, h)$.

Case 2: $n > n_c^C$. Under σ, this implies $m^t(\sigma, h) = 0$ and thus $y_f(h) = d$. Since $\hat{\sigma} \neq \sigma$ and $\hat{\sigma}$ agrees with σ at all histories $\tilde{h} \neq h$ and for all $\tilde{f} \neq f$, it must be that $\hat{y}_f(h) \neq y_f(h)$, and hence $\hat{y}_f(h) = c$. The expected present values are therefore

$$U_f^t(\sigma, h) = -R + (1 - p(n,0))\left(\frac{1}{n}\Pi_1 + \frac{n-1}{n}\Pi_0\right) + p(n,0)\,\beta\,U_f^{t+1}(\sigma, h)$$

and

$$U_f^t(\hat{\sigma}, h) = -R + (1 - p(n,1))\,\Pi_0 + p(n,1)\,\beta\,U_f^{t+1}(\hat{\sigma}, h).$$

Since $n > n_c^C$ and $n_c^C \geq 1$ (see Lemma A.21), we have $n \geq 2$, which implies $p(n, 0) = p(n, 1)$. Furthermore, since participation is fixed at $t = 0$, the number of active factions remains $n > n_c^C$ at all $\tilde{t} > t$ and $\tilde{h} \in H^{\tilde{t}}$ succeeding h. Since both strategies agree at all $\tilde{h} \neq h$, both prescribe $m^{\tilde{t}}(\sigma, \tilde{h}) = m^{\tilde{t}}(\hat{\sigma}, \tilde{h}) = 0$ at all such histories, so the continuation values coincide: $U_f^{t+1}(\sigma, h) = U_f^{t+1}(\hat{\sigma}, h)$. Together with $\Pi_1 > \Pi_0$, this yields $U_f^t(\sigma, h) > U_f^t(\hat{\sigma}, h)$. □

Lemma A.21. *Repression and $\tilde{n}_c^C$*
The cooperation threshold $\tilde{n}_c^C$ satisfies the following properties:

A.21.1 $\lim_{R\to 0} \tilde{n}_c^C \geq 1$.

A.21.2 $\lim_{R\to\infty} \tilde{n}_c^C = \infty$.

A.21.3 $\tilde{n}_c^C$ *is strictly increasing in* R.

Proof: Recall the closed-form solution of $\tilde{n}_c^C$ as defined in Lemma A.20.

A.21.1: At $R = 0$, the threshold satisfies

$$\tilde{n}_c^C(0) = \frac{\begin{array}{c}\Delta\Pi + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1) \\ + \sqrt{\Big(\Delta\Pi + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1)\Big)^2 + 4(1-\beta)\zeta_d \Delta\Pi^2}\end{array}}{2\Delta\Pi}.$$

Thus $\tilde{n}_c^C(0) \geq 1$ is equivalent to

$$\begin{aligned}&\Delta\Pi + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1) \\ &\quad + \sqrt{\Big(\Delta\Pi + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1)\Big)^2 + 4(1-\beta)\zeta_d \Delta\Pi^2} \;\geq\; 2\Delta\Pi.\end{aligned}$$

If $\Delta\Pi + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1) \geq 2\Delta\Pi$, the inequality is immediate. Otherwise, squaring both sides is valid and yields

$$(1-\beta)\zeta_d \Delta\Pi + \Delta\Pi + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1) \geq \Delta\Pi.$$

Using $\Delta\Pi = \Pi_1 - \Pi_0$, this reduces to

$$(1-\beta)\Pi_1(\zeta_d - \zeta_c) \geq 0,$$

which holds because $\beta \in (0,1)$, $\Pi_1 > 0$, and $\zeta_d > \zeta_c$. Hence $\lim_{R\to 0} \tilde{n}_c^C \geq 1$.

A.21.2: Since

$$\Delta\Pi + \beta R(\zeta_d - \zeta_c) + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1) \to \infty \quad \text{as } R \to \infty,$$

and $\beta(\zeta_d - \zeta_c) > 0$, the numerator of $\tilde{n}_c^C$ of Lemma A.20 diverges to infinity. Therefore, $\lim_{R\to\infty} \tilde{n}_c^C = \infty$.

A.21.3: Differentiating $\tilde{n}_c^C$ of Lemma A.20 with respect to R gives

$$\frac{\partial \tilde{n}_c^C}{\partial R} = \frac{\beta(\zeta_d - \zeta_c)}{2\Delta\Pi}\left[1 + \frac{\Delta\Pi + \beta R(\zeta_d - \zeta_c) + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1)}{\sqrt{\begin{array}{l}\Big(\Delta\Pi + \beta R(\zeta_d - \zeta_c) + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1)\Big)^2 \\ \quad + 4(1-\beta)\zeta_d \Delta\Pi^2\end{array}}}\right].$$

Since

$$\begin{aligned}&\sqrt{\Big(\Delta\Pi + \beta R(\zeta_d - \zeta_c) + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1)\Big)^2 + 4(1-\beta)\zeta_d \Delta\Pi^2} \\ &\qquad > \Big|\, \Delta\Pi + \beta R(\zeta_d - \zeta_c) + (1-\beta)(\zeta_d \Pi_0 - \zeta_c \Pi_1) \,\Big|,\end{aligned}$$

the bracketed factor in $\partial \tilde{n}_c^C / \partial R$ is strictly positive. Because $\zeta_d > \zeta_c$, it follows that $\partial \tilde{n}_c^C / \partial R > 0$, so $\tilde{n}_c^C$ is strictly increasing in R. $\square$

Proof of Lemma 7

Since n_c^C is the floor function of $\tilde{n}_c^C$, the lemma follows directly from Lemma A.21.

$\square$

Lemma A.22. ***Comparative statics of*** $\tilde{n}_c^C$

Let $V^* := V_c^C(\tilde{n}_c^C) = V_d^C(\tilde{n}_c^C)$ *and suppose* $V^* \geq 0$. *Further comparative statics of the cooperation threshold* $\tilde{n}_c^C$ *with respect to the remaining parameters are:*

A.22.1 $\tilde{n}_c^C$ *is strictly decreasing in* Π_1.

A.22.2 $\tilde{n}_c^C$ *is increasing in* Π_0, *and strictly increasing whenever* $\tilde{n}_c^C > 1$.

A.22.3 $\tilde{n}_c^C$ *is strictly decreasing in* ζ_c.

A.22.4 $\tilde{n}_c^C$ *is strictly increasing in* ζ_d.

A.22.5 $\tilde{n}_c^C$ *is decreasing in* β, *and strictly decreasing whenever* $V^* > 0$.

Proof: Recall that $\Delta V_c^C(n) = V_c^C(n) - V_d^C(n)$, where

$$V_c^C(n) = \frac{\Delta\Pi + n\Pi_0 - R(\zeta_c + n)}{\zeta_c(1-\beta) + n}, \qquad V_d^C(n) = \frac{n\Pi_1 - R(\zeta_d + n)}{\zeta_d(1-\beta) + n}.$$

The threshold $\tilde{n}_c^C$ solves $\Delta V_c^C(\tilde{n}_c^C) = 0$. By Part A.20.2 of Lemma A.20, $\partial \Delta V_c^C / \partial n < 0$ at $\tilde{n}_c^C$, so the implicit function theorem applied to $\Delta V_c^C(\tilde{n}_c^C) = 0$ gives, for any parameter θ,

$$\frac{\partial \tilde{n}_c^C}{\partial \theta} = -\frac{\partial \Delta V_c^C / \partial \theta}{\partial \Delta V_c^C / \partial n}\bigg|_{n = \tilde{n}_c^C},$$

so the sign of $\partial \tilde{n}_c^C / \partial \theta$ is the same as the sign of $\partial \Delta V_c^C / \partial \theta$ evaluated at the threshold. The denominators $\zeta_c(1-\beta) + n$ and $\zeta_d(1-\beta) + n$ are strictly positive.

A.22.1: Since $\Delta\Pi = \Pi_1 - \Pi_0$, the parameter Π_1 enters both value functions, and

$$\begin{aligned} \frac{\partial \Delta V_c^C}{\partial \Pi_1} &= \frac{1}{\zeta_c(1-\beta) + n} - \frac{n}{\zeta_d(1-\beta) + n} \\ &= \frac{N(n)}{\Big(\zeta_c(1-\beta) + n\Big)\Big(\zeta_d(1-\beta) + n\Big)}, \end{aligned}$$

where

$$N(n) := -n^2 + \Big(1 - (1-\beta)\zeta_c\Big)n + (1-\beta)\zeta_d.$$

The denominators are positive, so this derivative has the same sign as $N(\tilde{n}_c^C)$, which we now show is negative. The parabola N opens downward and satisfies $N(1) = (1-\beta)(\zeta_d - \zeta_c) > 0$, so it is positive up to its larger root n_N and negative beyond. Hence $N(\tilde{n}_c^C) < 0$ holds precisely when $\tilde{n}_c^C > n_N$, and it remains to establish this inequality. Let Q be the quadratic from (27), of which $\tilde{n}_c^C$ is the larger root. Since Q opens downward, it is positive only between its two roots, so any point at which Q is positive lies below $\tilde{n}_c^C$. Evaluating Q at n_N and using $n_N^2 = \big(1-(1-\beta)\zeta_c\big)n_N + (1-\beta)\zeta_d$ to eliminate n_N^2 gives

$$Q(n_N) = n_N\,(\zeta_d - \zeta_c)\,\Big((1-\beta)\Pi_0 + \beta R\Big) > 0,$$

since $n_N > 0$. Therefore $n_N < \tilde{n}_c^C$, so $N(\tilde{n}_c^C) < 0$, and $\tilde{n}_c^C$ is strictly decreasing in Π_1.

A.22.2: Writing $V_c^C(n) = \dfrac{\Pi_1 + (n-1)\Pi_0 - R(\zeta_c + n)}{\zeta_c(1-\beta) + n}$,

$$\frac{\partial \Delta V_c^C}{\partial \Pi_0} = \frac{n-1}{\zeta_c(1-\beta) + n}.$$

The denominator is positive, so the sign of this derivative is the same as the sign of $n-1$, which is non-negative at the threshold since $\tilde{n}_c^C \geq 1$ and strictly positive when $\tilde{n}_c^C > 1$. Hence $\tilde{n}_c^C$ is increasing in Π_0, and strictly increasing whenever $\tilde{n}_c^C > 1$.

A.22.3: Only V_c^C depends on ζ_c. Differentiating gives

$$\begin{aligned}
\frac{\partial \Delta V_c^C}{\partial \zeta_c} &= \frac{-R\Big(\zeta_c(1-\beta)+n\Big) - (1-\beta)\Big(\Delta\Pi + n\Pi_0 - R(\zeta_c+n)\Big)}{\Big(\zeta_c(1-\beta)+n\Big)^2} \\
&= -\frac{R + (1-\beta)V_c^C(n)}{\zeta_c(1-\beta)+n}.
\end{aligned}$$

At the threshold $V_c^C(\tilde{n}_c^C) = V^*$, so this equals $-\dfrac{R+(1-\beta)V^*}{\zeta_c(1-\beta)+\tilde{n}_c^C}$, which is strictly negative when $V^* \geq 0$, since then $R + (1-\beta)V^* > 0$. Hence $\tilde{n}_c^C$ is strictly decreasing in ζ_c.

A.22.4: Only V_d^C depends on ζ_d, entering ΔV_c^C with a negative sign. Differentiating gives

$$\frac{\partial \Delta V_c^C}{\partial \zeta_d} = \frac{R\Big(\zeta_d(1-\beta)+n\Big) + (1-\beta)\Big(n\Pi_1 - R(\zeta_d+n)\Big)}{\Big(\zeta_d(1-\beta)+n\Big)^2}$$
$$= \frac{R + (1-\beta)V_d^C(n)}{\zeta_d(1-\beta)+n}.$$

At the threshold $V_d^C(\tilde{n}_c^C) = V^*$, so this equals $\dfrac{R+(1-\beta)V^*}{\zeta_d(1-\beta)+\tilde{n}_c^C}$, which is strictly positive when $V^* \geq 0$. Hence $\tilde{n}_c^C$ is strictly increasing in ζ_d.

A.22.5: The parameter β enters both value functions. Differentiating gives

$$\frac{\partial \Delta V_c^C}{\partial \beta} = \frac{\zeta_c\Big(\Delta\Pi + n\Pi_0 - R(\zeta_c+n)\Big)}{\Big(\zeta_c(1-\beta)+n\Big)^2} - \frac{\zeta_d\Big(n\Pi_1 - R(\zeta_d+n)\Big)}{\Big(\zeta_d(1-\beta)+n\Big)^2}$$
$$= \frac{\zeta_c V_c^C(n)}{\zeta_c(1-\beta)+n} - \frac{\zeta_d V_d^C(n)}{\zeta_d(1-\beta)+n}.$$

At the threshold $V_c^C(\tilde{n}_c^C) = V_d^C(\tilde{n}_c^C) = V^*$, so this equals

$$\left.\frac{\partial \Delta V_c^C}{\partial \beta}\right|_{n=\tilde{n}_c^C} = V^* \frac{\tilde{n}_c^C(\zeta_c-\zeta_d)}{\Big(\zeta_c(1-\beta)+\tilde{n}_c^C\Big)\Big(\zeta_d(1-\beta)+\tilde{n}_c^C\Big)}.$$

Since $\tilde{n}_c^C(\zeta_c-\zeta_d) < 0$ and the denominator is positive, this is always non-positive if $V^* \geq 0$ and thus $\tilde{n}_c^C$ is decreasing in β. Moreover, the expression is strictly negative when $V^* > 0$. Hence $\tilde{n}_c^C$ is then strictly decreasing in β. $\square$

Lemma A.23. ***Comparative statics of*** n_c^C
Let $V^* := V_c^C(\tilde{n}_c^C) = V_d^C(\tilde{n}_c^C)$ *and suppose* $V^* \geq 0$. *Further comparative statics of* n_c^C *with respect to the remaining parameters are:*

A.23.1 n_c^C *is decreasing in* Π_1.

A.23.2 n_c^C *is increasing in* Π_0.

A.23.3 n_c^C *is decreasing in* ζ_c.

A.23.4 n_c^C *is increasing in* ζ_d.

A.23.5 n_c^C *is decreasing in* β.

Proof: Since n_c^C is the floor function of $\tilde{n}_c^C$, it follows, as noted above, that for any variable θ, n_c^C is increasing (decreasing) in θ whenever $\tilde{n}_c^C$ is increasing (decreasing) in θ. The result then follows directly from Lemma A.22, which establishes the corresponding comparative statics for $\tilde{n}_c^C$ with respect to Π_1, Π_0, ζ_c, ζ_d, and β. □

Lemma A.24. ***Properties of*** $\tilde{n}_d^C$

Define

$$\tilde{n}_d^C \geq 0 \quad \textit{such that} \quad V_d^D(\tilde{n}_d^C + 1) - V_i^C(\tilde{n}_d^C) = 0.$$

A.24.1 The threshold $\tilde{n}_d^C \geq 0$ *is well defined if and only if*

$$R \leq \frac{\Pi_1}{\zeta_d + 1}.$$

A.24.2 If $\tilde{n}_d^C \geq 0$ *exists, it is unique.*

A.24.3 If $\tilde{n}_d^C \geq 0$ *exists, then*

$$V_d^{D\prime}(\tilde{n}_d^C + 1) < V_i^{C\prime}(\tilde{n}_d^C).$$

Proof: Recall that

$$V_d^D(n+1) = \frac{\Delta\Pi - R\zeta_d + (n+1)(\Pi_0 - R)}{\zeta_d(1-\beta) + n + 1}, \qquad V_i^C(n) = \frac{n\Pi_0}{\zeta_c(1-\beta) + n}.$$

The threshold $\tilde{n}_d^C$ solves $V_d^D(\tilde{n}_d^C + 1) - V_i^C(\tilde{n}_d^C) = 0$. Clearing denominators gives $Q(\tilde{n}_d^C) = 0$, where

$$\begin{aligned} Q(n) := & -Rn^2 + \Big(\Delta\Pi - R(\zeta_d + 1) + \zeta_c(1-\beta)(\Pi_0 - R) - \Pi_0\zeta_d(1-\beta)\Big)n \\ & + \zeta_c(1-\beta)\Big(\Pi_1 - R(\zeta_d + 1)\Big). \end{aligned} \tag{28}$$

A.24.1: The constant term in (28), $\zeta_c(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)$, has the sign of $\Pi_1 - R(\zeta_d + 1)$. If $R \leq \frac{\Pi_1}{\zeta_d+1}$, then the constant term is non-negative. Since the coefficient on n^2 is $-R < 0$, the parabola opens downward and is non-negative at $n = 0$, and since $Q(n) \to -\infty$ as $n \to +\infty$, it must cross zero for some $n > 0$. Hence $\tilde{n}_d^C \geq 0$ exists. Conversely, suppose $R > \frac{\Pi_1}{\zeta_d+1}$. If no real roots exist, $\tilde{n}_d^C$ is not defined. Suppose real roots exist. The constant term is negative, so Vieta's formulas imply the product of the two roots is positive. The sum of the two roots equals

$$\frac{\Delta\Pi - R(\zeta_d + 1) + \zeta_c(1-\beta)(\Pi_0 - R) - \Pi_0\zeta_d(1-\beta)}{R}.$$

The numerator is strictly negative: $R > \frac{\Pi_1}{\zeta_d+1}$ implies $R(\zeta_d + 1) > \Pi_1 > \Delta\Pi$, and the remaining terms satisfy

$$\zeta_c(1-\beta)(\Pi_0 - R) - \Pi_0\zeta_d(1-\beta) = -\Big(R\zeta_c(1-\beta) + \Pi_0(1-\beta)(\zeta_d - \zeta_c)\Big) < 0,$$

since $\zeta_d > \zeta_c$. Hence the sum of the two roots is negative. A positive product and a negative sum imply both roots are negative, so no non-negative solution exists.

A.24.2: The product of the roots of (28) equals

$$\frac{\zeta_c(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)}{-R},$$

which is non-positive whenever $\tilde{n}_d^C \geq 0$ exists, by Part A.24.1. Two real roots with non-positive product have weakly opposite signs, so at most one can be non-negative. Hence $\tilde{n}_d^C \geq 0$, if it exists, is unique.

A.24.3: We first show that if $\tilde{n}_d^C \geq 0$ exists, then the quadratic in (28) has two distinct real roots. The discriminant is

$$\Big(\Delta\Pi - R(\zeta_d+1) + \zeta_c(1-\beta)(\Pi_0 - R) - \Pi_0\zeta_d(1-\beta)\Big)^2 + 4R\zeta_c(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big).$$

If $R < \frac{\Pi_1}{\zeta_d+1}$, then the second term is strictly positive, so the discriminant is positive. If $R = \frac{\Pi_1}{\zeta_d+1}$, then it equals

$$\Big(-\Pi_0(1 + (\zeta_d - \zeta_c)(1-\beta)) - R\zeta_c(1-\beta)\Big)^2 > 0.$$

Hence the discriminant is positive, and the quadratic has two distinct real roots. We rewrite

$$V_d^D(n) - V_i^C(n) = \frac{Q(n)}{(\zeta_d(1-\beta)+n+1)(\zeta_c(1-\beta)+n)},$$

where $Q(n)$ is the quadratic in (28). Since $Q(\tilde{n}_d^C) = 0$ and the denominator is strictly positive at $\tilde{n}_d^C$, it follows that

$$V_d^{D\prime}(\tilde{n}_d^C + 1) - V_i^{C\prime}(\tilde{n}_d^C) = \frac{Q'(\tilde{n}_d^C)}{(\zeta_d(1-\beta)+\tilde{n}_d^C+1)(\zeta_c(1-\beta)+\tilde{n}_d^C)}.$$

Since $Q(n)$ is a downward-opening parabola and $\tilde{n}_d^C$ is the larger of its two real roots, $Q'(\tilde{n}_d^C) < 0$, and thus $V_d^{D\prime}(\tilde{n}_d^C + 1) - V_i^{C\prime}(\tilde{n}_d^C) < 0$. □

Lemma A.25. ***Comparative statics of*** $\tilde{n}_d^C$
Suppose that $\tilde{n}_d^C > 0$ *of Lemma A.24 exists. Then the cooperation-breaking entry threshold satisfies the following comparative statics:*

A.25.1 $\tilde{n}_d^C$ *is strictly decreasing in* R.

A.25.2 $\tilde{n}_d^C < \tilde{n}^D$.

A.25.3 $\tilde{n}_d^C < \tilde{n}_a^C$.

Proof: Recall that

$$V_d^D(n+1) = \frac{\Delta\Pi - R\zeta_d + (n+1)(\Pi_0 - R)}{\zeta_d(1-\beta)+n+1}, \qquad V_i^C(n) = \frac{n\Pi_0}{\zeta_c(1-\beta)+n}.$$

The cooperation-breaking entry threshold $\tilde{n}_d^C$ solves

$$\Delta V^{DC}(\tilde{n}_d^C) = V_d^D(\tilde{n}_d^C+1) - V_i^C(\tilde{n}_d^C) = 0.$$

For any scalar parameter θ, the implicit function theorem gives

$$\frac{\partial \tilde{n}_d^C}{\partial \theta} = -\frac{\partial \Delta V^{DC}/\partial\theta}{\partial \Delta V^{DC}/\partial \tilde{n}_d^C}.$$

By Lemma A.24, $\partial\Delta V^{DC}/\partial\tilde{n}_d^C < 0$ at the equilibrium intersection. Hence the sign of $\partial\tilde{n}_d^C/\partial\theta$ is the same as the sign of $\partial\Delta V^{DC}/\partial\theta$.

A.25.1: Differentiating ΔV^{DC} with respect to R gives

$$\frac{\partial \Delta V^{DC}}{\partial R} = -\frac{\zeta_d + \tilde{n}_d^C + 1}{\zeta_d(1-\beta)+\tilde{n}_d^C+1} < 0,$$

and therefore $\partial\tilde{n}_d^C/\partial R < 0$.

A.25.2: First note that $\tilde{n}_d^C$ exists only if $R < \frac{\Pi_1}{\zeta_d+1}$, which implies that $\tilde{n}^D$ exists. Let

$$V_i(n;\zeta) = \frac{n\Pi_0}{\zeta(1-\beta)+n},$$

and let $n^*(\zeta) \geq 0$ denote the threshold solving

$$\Delta V(n^*;\zeta) = V_d^D(n^*+1) - V_i(n^*;\zeta) = 0.$$

By the implicit function theorem,

$$\frac{\partial n^*}{\partial \zeta} = -\frac{\partial \Delta V/\partial \zeta}{\partial \Delta V/\partial n},$$

where

$$\frac{\partial \Delta V}{\partial \zeta} = -\frac{\partial V_i}{\partial \zeta} = \frac{n^*\,\Pi_0(1-\beta)}{\big(\zeta(1-\beta)+n^*\big)^2} > 0$$

for $n^* > 0$, and $\partial \Delta V/\partial n < 0$ at the threshold, by the same argument as in Part A.24.3 of Lemma A.24. Hence $n^*(\zeta)$ is strictly increasing in ζ. Since $\tilde{n}^D = n^*(\zeta_d)$, $\tilde{n}^C_d = n^*(\zeta_c)$, and $\zeta_d > \zeta_c$, it follows that $\tilde{n}^C_d < \tilde{n}^D$.

A.25.3: First note that $\tilde{n}^C_d$ exists only if $R < \frac{\Pi_1}{\zeta_d+1}$, which implies that $R < \frac{\Pi_1}{\zeta_c+1}$ and thus the existence of $\tilde{n}^C_a$. Let

$$V_a(n;\zeta) = \frac{\Delta\Pi - \zeta R + (n+1)(\Pi_0 - R)}{\zeta(1-\beta)+n+1},$$

and let $n^*(\zeta) \geq 0$ denote the threshold solving

$$\Delta V(n^*;\zeta) = V_a(n^*+1;\zeta) - V_i^C(n^*) = 0.$$

By the implicit function theorem,

$$\frac{\partial n^*}{\partial \zeta} = -\frac{\partial \Delta V/\partial \zeta}{\partial \Delta V/\partial n},$$

where

$$\begin{aligned}\frac{\partial \Delta V}{\partial \zeta} &= \frac{\partial V_a}{\partial \zeta} \\ &= -\frac{\begin{array}{l} R\big(\zeta(1-\beta)+n^*+1\big) \\ \quad + (1-\beta)\big(\Delta\Pi - R\zeta + (n^*+1)(\Pi_0 - R)\big)\end{array}}{\big(\zeta(1-\beta)+n^*+1\big)^2} \;<\; 0.\end{aligned}$$

The inequality follows since $V_i^C(n^*) > 0$ and $\Delta V(n^*;\zeta) = 0$ imply $V_a(n^* + 1;\zeta) > 0$ and thus $\Delta\Pi - R\zeta + (n^*+1)(\Pi_0 - R) > 0$. Since $\partial V_a/\partial \zeta < 0$ and $\partial \Delta V/\partial n < 0$ at the threshold n^*, by the same argument as in Part A.24.3 of Lemma A.24, $n^*(\zeta)$ is strictly decreasing in ζ. Since $\tilde{n}^C_a = n^*(\zeta_c)$, $\tilde{n}^C_d = n^*(\zeta_d)$, and $\zeta_d > \zeta_c$, it follows that $\tilde{n}^C_d < \tilde{n}^C_a$. $\square$

Lemma A.26. ***Cooperation collapses above the threshold***
For any $t \geq 1$, $h \in H^t$, and opposition equilibrium σ,

$$n(\sigma, h) > n_c^C \Rightarrow m^t(\sigma, h) = 0.$$

Proof: Consider some opposition equilibrium σ and suppose by contradiction that at some $t \geq 1$ and $h \in H^t$ it holds that

$$n(\sigma, h) > n_c^C \text{ and } m^t(\sigma, h) > 0.$$

Consider any $f \in F_a(\sigma, h)$ such that $y_f(h) = c$. Such f must exist if $m^t(\sigma, h) > 0$. Consider $\hat{\sigma} \neq \sigma$ that differs from σ only with respect to the interaction behavior of f at history h. Hence, $\sigma_{\tilde{f}} = \hat{\sigma}_{\tilde{f}}$ for all $\tilde{f} \neq f$ and $\sigma_f(\tilde{h}) = \hat{\sigma}_f(\tilde{h})$ for all $\tilde{h} \neq h$. Since σ is an opposition equilibrium, the interaction behavior of all $\tilde{f}$ at periods $\tilde{t} > t$ does not condition on the behavior at t. Since σ and $\hat{\sigma}$ coincide at all histories $\tilde{h} \neq h$, they must coincide at all future $\tilde{t} > t$. Thus, the continuation utility of $\hat{\sigma}$ and σ coincide, that is, $U_f^{t+1}(\sigma, h) = U_f^{t+1}(\hat{\sigma}, h) =: \bar{U}$. For notational brevity, let $n(\sigma, h) = n(\hat{\sigma}, h) = n$, and note that the equality follows since the number of active factions is determined solely by history h. Moreover, let $m^t(\sigma, h) = m^t$ and $m^t(\hat{\sigma}, h) = \hat{m}^t$, and note that $\hat{m}^t = m^t - 1$, since only f changes its interaction behavior at h, from cooperating to defecting. For σ to be an opposition equilibrium, it must be a subgame perfect Nash equilibrium, implying it must yield a weakly larger expected present value than $\hat{\sigma}$:

$$U_f^t(\sigma, h) \geq U_f^t(\hat{\sigma}, h),$$

where

$$\begin{aligned} U_f^t(\sigma, h) = & -R + (1 - p(n, m^t))\Big(\hat{p}(a, n, c, m^t)\Pi_1 + (1 - \hat{p}(a, n, c, m^t))\Pi_0\Big) \\ & + p(n, m^t)\beta\bar{U} \end{aligned}$$

and

$$\begin{aligned} U_f^t(\hat{\sigma}, h) = & -R + (1 - p(n, \hat{m}^t))\Big(\hat{p}(a, n, d, \hat{m}^t)\Pi_1 + (1 - \hat{p}(a, n, d, \hat{m}^t))\Pi_0\Big) \\ & + p(n, \hat{m}^t)\beta\bar{U}. \end{aligned}$$

Case 1: $m^t \in (0, n)$. In this case, $\hat{p}(a, n, c, m^t) = 0$, $p(n, m^t) = p(n, \hat{m}^t) =$

$p(n, 0)$, and $\hat{p}(a, n, d, \hat{m}^t) > 0$. Substituting and rearranging yields

$$U_f^t(\sigma, h) \geq U_f^t(\hat{\sigma}, h) \Leftrightarrow \Pi_0 \geq \Pi_1.$$

Since $\Pi_0 \geq \Pi_1$ is not true, $U_f^t(\sigma, h) \geq U_f^t(\hat{\sigma}, h)$ cannot hold.

Case 2: $m^t = n$. Since σ is an opposition equilibrium, it must hold for all $\tilde{t} > t$ and histories $\tilde{h} \in H^{\tilde{t}}$ that coincide with h on the participation decision of factions that $m^t(\sigma, h) = m^{\tilde{t}}(\sigma, \tilde{h})$. Hence, for all $\tilde{t} > t$ and $\tilde{h} \in H^{\tilde{t}}$ that build on h, $m^{\tilde{t}}(\sigma, \tilde{h}) = n$. Since the path of play generated by σ from h is itself one such continuation $\tilde{h}$, it follows that $m^{\tilde{t}}(\sigma, h) = n$ for all future periods $\tilde{t} > t$. Thus, along the path of play there are n active factions who all cooperate, implying that the expected present value is $\bar{U} = V_c^C(n)$. Note that under σ this also holds at t, so that $U_f^t(\sigma, h) = V_c^C(n)$. Since $m^t = n$, it follows that $\hat{p}(a, n, c, m^t) = 1/n$, $p(n, m^t) > p(n, \hat{m}^t) = p(n, 0)$, and $\hat{p}(a, n, d, \hat{m}^t) = 1$. Thus,

$$U_f^t(\sigma, h) = V_c^C(n)$$

and

$$U_f^t(\hat{\sigma}, h) = -R + (1 - p(n, 0))\Pi_1 + p(n, 0)\beta V_c^C(n).$$

Substituting and rearranging yields

$$U_f^t(\sigma, h) - U_f^t(\hat{\sigma}, h) > 0 \Leftrightarrow V_c^C(n) > \frac{-R + (1 - p(n, 0))\Pi_1}{1 - \beta p(n, 0)} = V_d^C(n).$$

However, Lemma A.20 implies that $n > n_c^C$ gives $V_c^C(n) < V_d^C(n)$, so we reach a contradiction. Hence $m^t = n$ is also not possible, and thus the lemma holds. $\square$

Lemma A.27. ***Repression and*** $\bar{R}^C$
There exists $\bar{R}^C \in [0, \frac{\Pi_1}{\zeta_c+1})$ *such that*

$$R \geq \bar{R}^C \Leftrightarrow n_c^C \geq \tilde{n}_d^C.$$

Proof: To prove the lemma, we construct $\bar{R}^C$ that satisfies the above condition. Recall that $n_c^C = \max\{x \in \mathbb{N} : x \leq \tilde{n}_c^C\}$, where $\tilde{n}_c^C(R)$ is strictly increasing from 1 to ∞ on some domain (see Lemma A.21). Hence, n_c^C is a increasing step function in R, taking all integer values from 1 to ∞. Consider $\tilde{n}_d^C$ as defined in Lemma A.24. Here $\tilde{n}_d^C \to 0$ as $R \to \frac{\Pi_1}{\zeta_d+1}$, since at this value $V_d^D(n+1) = V_i^C(n)$ holds at $n = 0$ and such a nonnegative crossing must be unique. Furthermore, recall that Lemma A.25 establishes that $\tilde{n}_d^C$ is strictly decreasing in R.

Case 1: $\lim_{R\to 0} \tilde{n}_d^C > \lim_{R\to 0} n_c^C$. Recall that n_c^C is increasing in R and always above 1, and $\tilde{n}_d^C$ strictly decreases in R with $\lim_{R\to\frac{\Pi_1}{\zeta_d+1}} \tilde{n}_d^C = 0 < 1 \leq \lim_{R\to\frac{\Pi_1}{\zeta_d+1}} n_c^C$. These properties imply the existence of a unique quasi-crossing point $\bar{R}^C > 0$ where $R < \bar{R}^C \Rightarrow n_c^C < \tilde{n}_d^C$ and $R > \bar{R}^C \Rightarrow n_c^C > \tilde{n}_d^C$. As R approaches $\bar{R}^C$ from above, it must hold that $n_c^C > \tilde{n}_d^C$. Since n_c^C is right-continuous and $\tilde{n}_d^C$ is continuous, the inequality holds weakly in the limit, so at $R = \bar{R}^C$ we have $n_c^C \geq \tilde{n}_d^C$. Thus,

$$R \geq \bar{R}^C \;\Leftrightarrow\; n_c^C \geq \tilde{n}_d^C.$$

Case 2: $\lim_{R\to 0} \tilde{n}_d^C \leq \lim_{R\to 0} n_c^C$. Since $\tilde{n}_d^C$ strictly decreases and n_c^C increases in R, it follows that $\tilde{n}_d^C < n_c^C$ always holds for $R > 0$. Thus, for $\bar{R}^C = 0$, the lemma holds. $\square$

Proof of Proposition 2

Throughout the following proof, let $\tilde{n}_d^C$ be defined as in Lemma A.24, and let $\bar{R}^C \in [0, \frac{\Pi_1}{\zeta_c+1})$ be such that

$$R \geq \bar{R}^C \Leftrightarrow n_c^C \geq \tilde{n}_d^C.$$

Such $\bar{R}^C$ exists by Lemma A.27.

"If": We first show that if $R \in [\bar{R}^C, \frac{\Pi_1}{\zeta_c+1}]$, then there is an opposition equilibrium as described. We do so by providing an explicit example. Suppose $R \in [\bar{R}^C, \frac{\Pi_1}{\zeta_c+1}]$. Let $F^C \subseteq F$ with $|F^C| = n^C$ and define $\sigma = (\sigma_f)_{f \in F}$ with $\sigma_f = (x_f, y_f)$ as follows:

1. For all $f \in F$:
$$x_f = a \Leftrightarrow f \in F^C$$

2. For all $t \geq 1$, $h \in H^t$, and $f \in F_a(\sigma, h)$:
   (a) $n(\sigma, h) \leq n_c^C \Rightarrow y_f(h) = c$ and
   (b) $n(\sigma, h) > n_c^C \Rightarrow y_f(h) = d$.

By construction, $n(\sigma, \emptyset) = n^C$, which follows from Condition 1 and $|F^C| = n^C$. Since $n^C \leq n_c^C$, we have $n(\sigma, h) \leq n_c^C$ at every history h on the path of play. Condition 2 then implies $y_f(h) = c$ for all active f and all on-path histories h, and hence $m^t(\sigma, \emptyset) = n^C$ for all $t \geq 1$. Hence, $\sigma \in \{\sigma : n(\sigma, \emptyset) = m^t(\sigma, \emptyset) = n^C \text{ for all } t \geq 1\}$.

We continue to show that σ is an opposition equilibrium. Since x_f is constant and $y_f(h)$ depends only on $n(\sigma, h)$, which is itself determined by $\{x_{\hat{f}}\}_{\hat{f} \in F}$, Condition 2 of Definition 1 holds. It remains to show that σ is a subgame perfect Nash equilibrium, for which we employ the one-shot deviation principle.

First, we investigate whether any faction has an incentive to deviate on the interaction dimension. Consider any $t \geq 1$, $h \in H^t$, and $f \in F_a(\sigma, h)$. By construction of σ,

$$n(\sigma, h) \leq n_c^C \Rightarrow y_{\tilde{f}}(h) = c \; \forall \tilde{f} \in F_a(\sigma, h) \Rightarrow m^t(\sigma, h) = n(\sigma, h)$$

and

$$n(\sigma, h) > n_c^C \Rightarrow y_{\tilde{f}}(h) = d \; \forall \tilde{f} \in F_a(\sigma, h) \Rightarrow m^t(\sigma, h) = 0.$$

Lemma 6 then implies that no faction f has a profitable one-shot deviation on the interaction dimension at (t, h). Since t and h were arbitrary, this holds for all $t \geq 1$ and $h \in H^t$.

Second, we show that no faction has an incentive to deviate on the participation dimension. Consider any $f \in F$ and let $\hat{\sigma}$ be such that $\hat{\sigma}_{\tilde{f}} = \sigma_{\tilde{f}}$ for all $\tilde{f} \neq f$, with $\hat{\sigma}_f = (\hat{x}_f, y_f)$, that is, $\hat{\sigma}$ differs from σ only in the participation decision of f.

Case 1: $n_a^C < n_c^C$, or $n_a^C = n_c^C$ and $x_f = a$. In both cases, $n^C = n_a^C \leq n_c^C$ and thus $n(\sigma, \emptyset) = n_a^C \leq n_c^C$. Since only f changes its participation decision under $\hat{\sigma}$ as compared to σ, it must hold that $n(\hat{\sigma}, \emptyset) \in \{n(\sigma, \emptyset) - 1, n(\sigma, \emptyset) + 1\}$. If $n_a^C < n_c^C$, this implies $n(\hat{\sigma}, \emptyset) \leq n(\sigma, \emptyset) + 1 \leq n_c^C$. If $n_a^C = n_c^C$ and $x_f = a$, the deviation is from $x_f = a$ to $\hat{x}_f = i$, implying $n(\hat{\sigma}, \emptyset) \leq n(\sigma, \emptyset) - 1 < n_c^C$. Thus, in both cases $n(\hat{\sigma}, \emptyset) \leq n_c^C$. By construction of σ and since $\hat{\sigma}$ and σ coincide on $y_{\tilde{f}}$ for all $\tilde{f}$, it follows that $y_{\tilde{f}}(h) = c$ for all $\tilde{f} \in F_a(\hat{\sigma}, h)$ and all histories h along the path of play. Consequently, $m^t(\hat{\sigma}, \emptyset) = n(\hat{\sigma}, \emptyset)$ for all $t \geq 1$. Lemma 4 then implies $U_f^0(\hat{\sigma}, \emptyset) \leq U_f^0(\sigma, \emptyset)$.

Case 2: $n_a^C > n_c^C$, or $n_c^C = n_a^C$ and $x_f = i$. Consequently, $n^C = n_c^C$.

Case 2a: $x_f = a$ and $\hat{x}_f = i$. Since only f changes its participation decision under $\hat{\sigma}$ as compared to σ, and f is inactive under $\hat{\sigma}$ and active under σ, it must hold that $n(\hat{\sigma}, \emptyset) = n(\sigma, \emptyset) - 1 = n_c^C - 1 < n_c^C$. By construction of σ and since $\hat{\sigma}$ and σ coincide on $y_{\tilde{f}}$ for all $\tilde{f}$, it follows that $y_{\tilde{f}}(h) = c$ for all $\tilde{f} \in F_a(\hat{\sigma}, h)$ and all histories h along the path of play. Consequently, $m^t(\hat{\sigma}, \emptyset) = n(\hat{\sigma}, \emptyset)$ for all $t \geq 1$. Thus, under $\hat{\sigma}$, f is inactive and there are $n_c^C - 1$ factions who are active and cooperate in every period, so the resulting expected present value is $U_f^0(\hat{\sigma}, \emptyset) = V_i^C(n_c^C - 1)$. Under σ, f is active and there are n_c^C factions who are active and cooperate in every period, so the resulting expected present value is $U_f^0(\sigma, \emptyset) = V_c^C(n_c^C)$. Recall that Lemma A.14 implies $V_c^C(n+1) \geq V_i^C(n) \Leftrightarrow n \leq \tilde{n}_a^C$. Thus, $V_c^C(n) \geq V_i^C(n-1)$ for all $n \leq n_a^C$. Since $n_c^C \leq n_a^C$, it holds that $V_c^C(n_c^C) \geq V_i^C(n_c^C - 1)$, and thus $U_f^0(\sigma, \emptyset) \geq U_f^0(\hat{\sigma}, \emptyset)$.

Case 2b: $x_f = i$ and $\hat{x}_f = a$. Since only f changes its participation decision under $\hat{\sigma}$ as compared to σ, and f is active under $\hat{\sigma}$ and inactive under σ, it must hold that $n(\hat{\sigma}, \emptyset) = n(\sigma, \emptyset) + 1 = n_c^C + 1 > n_c^C$. By construction of σ and since $\hat{\sigma}$ and σ coincide on $y_{\tilde{f}}$ for all $\tilde{f}$, it follows that $y_{\tilde{f}}(h) = d$ for all $\tilde{f} \in F_a(\hat{\sigma}, h)$ and all histories h along the path of play. Consequently, $m^t(\hat{\sigma}, \emptyset) = 0$ for all $t \geq 1$. Thus, under $\hat{\sigma}$, f is active and there are $n_c^C + 1$

factions who are active and defect in every period, so the resulting expected present value is $U_f^0(\hat{\sigma}, \emptyset) = V_d^D(n_c^C + 1)$. Under σ, f is inactive and there are n_c^C factions who are active and cooperate in every period, so the resulting expected present value is $U_f^0(\sigma, \emptyset) = V_i^C(n_c^C)$. Recall that $R \geq \bar{R}^C$ implies $n_c^C \geq \tilde{n}_d^C$. Moreover, Lemma A.24 implies that for all $n \geq \tilde{n}_d^C$, $V_d^D(n+1) \leq V_i^C(n)$. Thus, $V_d^D(n_c^C + 1) \leq V_i^C(n_c^C)$, and thus $U_f^0(\hat{\sigma}, \emptyset) \leq U_f^0(\sigma, \emptyset)$.

"Only if": Next, we show that no opposition equilibrium exists if $R \notin \left[\bar{R}^C, \frac{\Pi_1}{\zeta_c+1}\right]$. First, suppose $R > \frac{\Pi_1}{\zeta_c+1}$. In that case, $\tilde{n}_a^C$ is not well-defined (recall Lemma A.14). Since n_a^C is defined in terms of $\tilde{n}_a^C$, it follows that n_a^C is not well-defined, and consequently n^C is not well-defined. Hence, there does not exist a value n^C such that an opposition equilibrium σ could satisfy $n(\sigma, \emptyset) = n^C$.

Suppose $R < \bar{R}^C$. By contradiction, suppose there is an opposition equilibrium σ such that $n(\sigma, \emptyset) = m^t(\sigma, \emptyset) = n^C$ for all $t \geq 1$. Note that $R < \bar{R}^C$ implies $n_c^C < \tilde{n}_d^C$ by definition of $\bar{R}^C$ above. Moreover, since $\tilde{n}_d^C < \tilde{n}_a^C$ by Lemma A.25, $R < \bar{R}^C$ implies $n_c^C < \tilde{n}_d^C \leq \tilde{n}_a^C \leq n_a^C$. Thus, $n^C = n_c^C$. Consider any faction f that is inactive under σ, that is, $x_f = i$. Then faction f's expected present value is $U_f^0(\sigma, \emptyset) = V_i^C(n_c^C)$. Consider any other strategy profile $\hat{\sigma}$ that differs from σ only with respect to the participation decision of faction f, that is, $\hat{x}_f = a$. Consequently, $n(\hat{\sigma}, \emptyset) = n(\sigma, \emptyset) + 1 = n_c^C + 1$. Since σ is an opposition equilibrium, it must be that for any $\tilde{f}$ the interaction choice $y_{\tilde{f}}(h)$ conditions only on the set of active factions at h. Since $\hat{\sigma}$ coincides with σ on the interaction decision of all $\tilde{f}$, it must be that $m^t(\hat{\sigma}, \emptyset) = m^{\check{t}}(\hat{\sigma}, \emptyset)$ for all t and $\check{t}$. Under $\hat{\sigma}$, the number of cooperating factions must be equal across all periods, since otherwise $\hat{\sigma}$ would condition on something other than the participation behavior. If that were the case, σ would do so as well and could not be an opposition equilibrium. Moreover, Lemma A.26 implies that for σ, and any $t \geq 1$ and $h \in H^t$, $n(\sigma, h) > n_c^C \Rightarrow m^t(\sigma, h) = 0$. Since $\hat{\sigma}$ and σ coincide on the interaction behavior of all $\tilde{f}$, it holds that for any $t \geq 1$ and $h \in H^t$, $n(\hat{\sigma}, h) > n_c^C \Rightarrow m^t(\hat{\sigma}, h) = 0$. Since $n(\hat{\sigma}, \emptyset) = n_c^C + 1 > n_c^C$, and thus $n(\hat{\sigma}, h) > n_c^C$ for all histories h along the path of play, it holds that $m^t(\hat{\sigma}, h) = 0$ for all $t \geq 1$ and h along the path of play, and thus $m^t(\hat{\sigma}, \emptyset) = 0$ for all $t \geq 1$. Faction f is active under $\hat{\sigma}$, when $n_c^C + 1$ factions are active in total and all defect. The corresponding expected present value at $t = 0$ is $U_f^0(\hat{\sigma}, \emptyset) = V_d^D(n_c^C + 1)$. Lemma A.24 implies that the unique positive crossing of $V_d^D(n+1) - V_i^C(n)$ with 0 occurs at $\tilde{n}_d^C$, and that at this crossing $V_d^{D\prime}(n+1) < V_i^{C\prime}(n)$. Hence, $n < \tilde{n}_d^C \Leftrightarrow V_d^D(n+1) > V_i^C(n)$. Moreover, by construction of $\bar{R}^C$, $R < \bar{R}^C$

implies $n_c^C < \tilde{n}_d^C$, and thus $V_d^D(n_c^C + 1) > V_i^C(n_c^C)$ and $U_f^0(\hat{\sigma}, \emptyset) > U_f^0(\sigma, \emptyset)$. Hence, faction f benefits by deviating from σ to $\hat{\sigma}$, implying that σ is not a subgame perfect Nash equilibrium. We reach a contradiction, implying that there is no opposition equilibrium σ if $R < \bar{R}^C$. $\square$

Proof of Lemma 8

Recall from the proof of Proposition 2 that for $R = \bar{R}^C$ it holds that $0 < n_c^C < n_a^C$. Moreover, at $R = \frac{\Pi_1}{\zeta_c+1}$, $\tilde{n}_a^C = 0$ solves the quadratic expression that defines $\tilde{n}_a^C$ in the proof of Lemma A.14. Since that lemma guarantees a unique non-negative root, it follows that $\tilde{n}_a^C = 0$ and thus, by definition, $n_a^C = 1$. Lemma 5 implies that n_a^C is decreasing in R. Moreover, Lemma A.21 implies that $n_c^C \geq 1$ for all R and is increasing in R.

Case 1: $n_c^C > 1$ at $R = \frac{\Pi_1}{\zeta_c+1}$. Define $f(R) := n_c^C(R) - n_a^C(R)$. Since n_c^C is increasing and n_a^C is decreasing, f is increasing on $[\bar{R}^C, \frac{\Pi_1}{\zeta_c+1}]$. We have $f(\bar{R}^C) < 0$, since $n_c^C < n_a^C$ at $R = \bar{R}^C$, and $f(\frac{\Pi_1}{\zeta_c+1}) > 0$, since $n_a^C = 1 < n_c^C$ at $R = \frac{\Pi_1}{\zeta_c+1}$ by assumption. Hence there exists $\check{R} \in (\bar{R}^C, \frac{\Pi_1}{\zeta_c+1})$ such that $n_c^C \leq n_a^C$ for all $R \in (\bar{R}^C, \check{R})$ and $n_c^C \geq n_a^C$ for all $R \in (\check{R}, \frac{\Pi_1}{\zeta_c+1})$. Since $n^C = \min\{n_c^C, n_a^C\}$, it follows that $n^C = n_c^C$ for $R \in (\bar{R}^C, \check{R})$, which is increasing, and $n^C = n_a^C$ for $R \in (\check{R}, \frac{\Pi_1}{\zeta_c+1})$, which is decreasing.

Case 2: $n_c^C = 1$ at $R = \frac{\Pi_1}{\zeta_c+1}$. Since $n_c^C \geq 1$ for all R and is increasing, $n_c^C = 1$ on the entire interval $[\bar{R}^C, \frac{\Pi_1}{\zeta_c+1}]$. Since n_a^C is decreasing and $n_a^C = 1$ at $R = \frac{\Pi_1}{\zeta_c+1}$, we also have $n_a^C \geq 1$ for all R on this interval. Hence $n^C = \min\{n_c^C, n_a^C\} = 1$ throughout $[\bar{R}^C, \frac{\Pi_1}{\zeta_c+1}]$. That is, n^C is constant on this interval, and thus both increasing and decreasing on it. $\square$

A.3 Formal analysis of the incumbent's choice of repression

Lemma A.28. ***The defection equilibrium prevails at low repression***
There exists $\tilde{R} > 0$ *such that*

$$R < \tilde{R} \Rightarrow p(n^D, 0) < p(n^C, n^C).$$

Proof: The inequality $p(n^D, 0) < p(n^C, n^C)$ is equivalent to $\frac{\zeta_d}{n^D+\zeta_d} < \frac{\zeta_c}{n^C+\zeta_c}$, which in turn is equivalent to $\zeta_d\, n^C < \zeta_c\, n^D$. Consider $\check{R}$ as in Lemma 8, and suppose that $R < \check{R}$. Then $n^C = n_c^C$, and thus $p(n^D, 0) < p(n^C, n^C)$ if $\zeta_d\, n_c^C < \zeta_c\, n^D$. Recall that the threshold $\tilde{n}^D$ solves (24) and is thus given by

$$\tilde{n}^D = \frac{\Delta\Pi - R(\zeta_d(2-\beta)+1) + \sqrt{\begin{aligned}&\Big(\Delta\Pi - R(\zeta_d(2-\beta)+1)\Big)^2 \\ &\quad + 4R\zeta_d(1-\beta)\Big(\Pi_1 - R(\zeta_d+1)\Big)\end{aligned}}}{2R}.$$

It follows that $\lim_{R\to 0} \tilde{n}^D = \infty$. Moreover, by definition of n^D, it must hold that $n^D > \tilde{n}^D$, and thus $\lim_{R\to 0} n^D = \infty$. In addition, $n_c^C \leq \tilde{n}_c^C$ by definition of n_c^C. Recall that the threshold $\tilde{n}_c^C$ is the larger solution to (27), and thus

$$\tilde{n}_c^C = \frac{\begin{aligned}&\Pi_0(\zeta_d(1-\beta)-1) - \Pi_1(\zeta_c(1-\beta)-1) + \beta R(\zeta_d - \zeta_c) \\ &+ \sqrt{\begin{aligned}&\Big(\Pi_0(\zeta_d(1-\beta)-1) - \Pi_1(\zeta_c(1-\beta)-1) + \beta R(\zeta_d-\zeta_c)\Big)^2 \\ &\quad + 4(1-\beta)\zeta_d\Delta\Pi^2\end{aligned}}\end{aligned}}{2\Delta\Pi}.$$

It is straightforward that $\lim_{R\to 0} \tilde{n}_c^C$ is finite, and thus $\lim_{R\to 0} n_c^C < \infty$. Since $\lim_{R\to 0} n_c^C < \infty$ and $\lim_{R\to 0} n^D = \infty$, there exists $\tilde{R} \leq \check{R}$ such that $R < \tilde{R}$ implies $\zeta_c n^D > \zeta_d n_c^C$. Since $\tilde{R} \leq \check{R}$ ensures $n^C = n_c^C$ on this range, it follows that $\zeta_c n^D > \zeta_d n^C$, and thus $p(n^D, 0) < p(n^C, n^C)$. □

Lemma A.29. ***A divergence bound for large ζ_d and small ζ_c***
For any real number $r \in \mathbb{R}$, if ζ_d is sufficiently large and ζ_c sufficiently small, then

$$(\zeta_d - \zeta_c - 1) + \sqrt{(\zeta_c + 2)^2 + 4(\zeta_d - \zeta_c)} - \sqrt{(\zeta_d + 1)^2 + 4(\zeta_d - \zeta_c)} > r.$$

Proof: Fix $r \in \mathbb{R}$. We show that the inequality holds whenever

$$\zeta_c \le 1 \qquad \text{and} \qquad \zeta_d \ge \max\left\{2,\ 1 + \left(\tfrac{r+5}{2}\right)^2\right\}.$$

Suppose these hold. Then in particular $\zeta_d - \zeta_c \ge 1 > 0$, so all square roots below are real. We bound the three terms of the left side separately.

First, since $\zeta_c \le 1$,

$$\zeta_d - \zeta_c - 1 \ \ge\ \zeta_d - 2.$$

Second, dropping the nonnegative term $(\zeta_c + 2)^2$ and using $\zeta_c \le 1$,

$$\sqrt{(\zeta_c + 2)^2 + 4(\zeta_d - \zeta_c)} \ \ge\ \sqrt{4(\zeta_d - \zeta_c)} \ \ge\ 2\sqrt{\zeta_d - 1}.$$

Third, since $\zeta_c > 0$ implies $4(\zeta_d - \zeta_c) < 4(\zeta_d + 1) + 4$,

$$\sqrt{(\zeta_d + 1)^2 + 4(\zeta_d - \zeta_c)} \ <\ \sqrt{(\zeta_d + 1)^2 + 4(\zeta_d + 1) + 4} \ =\ \sqrt{(\zeta_d + 3)^2} \ =\ \zeta_d + 3.$$

Adding the first two bounds and subtracting the third,

$$\begin{aligned}(\zeta_d - \zeta_c - 1) + \sqrt{(\zeta_c + 2)^2 + 4(\zeta_d - \zeta_c)} &- \sqrt{(\zeta_d + 1)^2 + 4(\zeta_d - \zeta_c)} \\ &> \ (\zeta_d - 2) + 2\sqrt{\zeta_d - 1} - (\zeta_d + 3) \\ &= \ 2\sqrt{\zeta_d - 1} - 5.\end{aligned}$$

If $r + 5 > 0$, then $\zeta_d \ge 1 + \left(\frac{r+5}{2}\right)^2$ gives

$$\sqrt{\zeta_d - 1} \ \ge\ \tfrac{r+5}{2}, \qquad \text{hence} \qquad 2\sqrt{\zeta_d - 1} - 5 \ \ge\ r.$$

If instead $r + 5 \le 0$, the conclusion is immediate from $2\sqrt{\zeta_d - 1} - 5 \ge -5 \ge r$. In either case the left side strictly exceeds r, which proves the lemma. $\square$

Lemma A.30. ***Repression can decrease*** P_I
If

$$\frac{(\zeta_d - \zeta_c - 1) + \sqrt{(\zeta_c+2)^2 + 4(\zeta_d - \zeta_c)} - \sqrt{(\zeta_d+1)^2 + 4(\zeta_d - \zeta_c)}}{2} > 1$$

and β *is sufficiently close to* 1*, then there exists* $R_1 < R_2$ *such that* $P_I(R_1) > P_I(R_2)$*.*

Proof: Consider $\tilde{n}_c^C$, $\tilde{n}_a^C$, and $\tilde{n}^D$, and note that each is a continuous function of R and β (on the domain on which it exists). Recall that $\tilde{n}_c^C$ is strictly increasing in R with a finite value at $R = 0$ (see Lemma A.21), while $\tilde{n}_a^C$ and $\tilde{n}^D$ are both strictly decreasing in R (see Lemmas A.15 and A.5) with $\tilde{n}_a^C \to \infty$ and $\tilde{n}^D \to \infty$ as $R \to 0$ (see Lemmas A.18 and A.8). Moreover, $\tilde{n}_a^C \to 0$ as $R \to \frac{\Pi_1}{\zeta_c+1}$, and $\tilde{n}^D \to 0$ as $R \to \frac{\Pi_1}{\zeta_d+1}$. It follows that (1) $\tilde{n}^D + 1$ and $\tilde{n}_c^C$ intersect exactly once on $\left(0, \frac{\Pi_1}{\zeta_d+1}\right)$, at a point we call R_1, and (2) $\tilde{n}_a^C$ and $\tilde{n}_c^C$ intersect exactly once on $\left(0, \frac{\Pi_1}{\zeta_c+1}\right)$, at a point we call R_2.

By construction, for any $R \geq R_1$ we have $\tilde{n}^D + 1 \leq \tilde{n}_c^C$, and hence $n^D \leq n_c^C$, and for any $R \leq R_2$ we have $\tilde{n}_a^C \geq \tilde{n}_c^C$, so that $n_a^C \geq n_c^C$ and thus $n^C = n_c^C$.

We first argue that $R_1 \geq \bar{R}^C$. Recall from Lemma A.25 that $\tilde{n}_d^C \leq \tilde{n}^D$. Since, for all $R \geq R_1$, $\tilde{n}^D + 1 \leq \tilde{n}_c^C$, it follows that $\tilde{n}_d^C < \tilde{n}_c^C$ for all $R \geq R_1$. Recall further from the proof of Proposition 2 that $\bar{R}^C$ satisfies $R \geq \bar{R}^C \iff n_c^C \geq \tilde{n}_d^C$. Hence $R \geq R_1 \implies \tilde{n}_d^C \leq \tilde{n}_c^C \implies R \geq \bar{R}^C$, and thus $R_1 \geq \bar{R}^C$.

For the remainder of the argument, suppose $R_1 < R_2$; we verify below that this holds under the conditions of the lemma. For all $R \in [R_1, R_2]$ we then have (a) $n^C = n_c^C$, and (b) $n_c^C \geq n^D$ whenever $R \leq \bar{R}^D$. Two cases follow. If $R \leq \bar{R}^D$, then $n^C \geq n^D$ together with $\zeta_d > \zeta_c$ gives $\frac{\zeta_d}{\zeta_d+n^D} > \frac{\zeta_c}{\zeta_c+n^C}$, so $P_I(R) = p(n^C, n^C)$. If instead $R > \bar{R}^D$, the defection equilibrium fails to exist; but since $[R_1, R_2] \subseteq \left[\bar{R}^C, \frac{\Pi_1}{\zeta_c+1}\right]$, the cooperation equilibrium does exist, so again $P_I(R) = p(n^C, n^C)$. Hence $P_I(R) = p(n^C, n^C)$ with $n^C = n_c^C$ throughout $[R_1, R_2]$.

Viewing n_c^C as a function of R, we then have $P_I(R_1) > P_I(R_2)$ whenever $n_c^C(R_1) < n_c^C(R_2)$, which in turn holds if $1 < \tilde{n}_c^C(R_2) - \tilde{n}_c^C(R_1)$ (the strict inequality guarantees an integer gap between the two).

We verify this in the limit $\beta \to 1$. There, R_1 solves $\tilde{n}_c^C(R_1) = \tilde{n}^D(R_1) + 1$, where $\lim_{\beta\to 1} \tilde{n}_c^C(R) = 1 + R\frac{\zeta_d - \zeta_c}{\Delta\Pi}$ and $\lim_{\beta \to 1}\left(\tilde{n}^D(R) + 1\right) = \frac{\Delta\Pi}{R} - \zeta_d$. Equating the two and rearranging gives

$$R^2\,\frac{\zeta_d-\zeta_c}{\Delta\Pi}+R(\zeta_d+1)-\Delta\Pi=0,$$

a quadratic with roots

$$R^0=\frac{\Delta\Pi}{\zeta_d-\zeta_c}\cdot\frac{-\zeta_d-1\pm\sqrt{(\zeta_d+1)^2+4(\zeta_d-\zeta_c)}}{2}.$$

Since R_1 is the unique non-negative root, it corresponds to the $+$ solution:

$$R_1=\frac{\Delta\Pi}{\zeta_d-\zeta_c}\cdot\frac{-\zeta_d-1+\sqrt{(\zeta_d+1)^2+4(\zeta_d-\zeta_c)}}{2}.$$

An analogous argument yields, in the same limit,

$$R_2=\frac{\Delta\Pi}{\zeta_d-\zeta_c}\cdot\frac{-\zeta_c-2+\sqrt{(\zeta_c+2)^2+4(\zeta_d-\zeta_c)}}{2}.$$

Since $\tilde{n}_c^C(R)=1+R\frac{\zeta_d-\zeta_c}{\Delta\Pi}$ in this limit,

$$\tilde{n}_c^C(R_2)-\tilde{n}_c^C(R_1)=(R_2-R_1)\,\frac{\zeta_d-\zeta_c}{\Delta\Pi},$$

and substituting the expressions for R_1 and R_2 gives

$$\begin{aligned}\tilde{n}_c^C(R_2)-\tilde{n}_c^C(R_1)=\frac{1}{2}\Big((\zeta_d-\zeta_c-1)&+\sqrt{(\zeta_c+2)^2+4(\zeta_d-\zeta_c)}\\&-\sqrt{(\zeta_d+1)^2+4(\zeta_d-\zeta_c)}\Big),\end{aligned}$$

which is greater than 1 by the assumption of the lemma. Since n_c^C is the floor of $\tilde{n}_c^C$, a gap strictly greater than 1 in the continuous threshold forces the floor to rise by at least one step. It follows that, in the limit $\beta\to 1$,

$$\tilde{n}_c^C(R_2)-\tilde{n}_c^C(R_1)>1\;\Rightarrow\;n_c^C(R_2)\geq n_c^C(R_1)+1,$$

and hence $P_I(R_1)>P_I(R_2)$.

Finally, since all functions involved are continuous in β, and $R_1(\beta)$, $R_2(\beta)$ vary continuously in β as roots of a continuous family of equations, the quantity $\tilde{n}_c^C(R_2(\beta))-\tilde{n}_c^C(R_1(\beta))$ is itself continuous in β. As it strictly exceeds 1 at $\beta=1$, it continues to do so in a neighborhood of $\beta=1$, so the result established in the limit extends to β sufficiently close to 1. This proves the lemma. $\square$

Proof of Proposition 3

Consider ζ_d sufficiently large and ζ_c sufficiently small for

$$\frac{(\zeta_d - \zeta_c - 1) + \sqrt{(\zeta_c + 2)^2 + 4(\zeta_d - \zeta_c)} - \sqrt{(\zeta_d + 1)^2 + 4(\zeta_d - \zeta_c)}}{2} > 1$$

to hold, which is possible by Lemma A.29. Lemma A.30 then implies that for sufficiently large β, there exists $R_1 < R_2$ such that $P_I(R_1) > P_I(R_2)$. Consider any such β. It remains to be shown that there also exist $R_3 < R_4$ such that $P_I(R_3) < P_I(R_4)$.

Consider $R_4 = \frac{\Pi_1}{\zeta_c+1}$. At R_4, the cooperation equilibrium exists. Moreover, $\tilde{n}_a^C = 0$ and thus $n_a^C = 1$. Since $n_c^C \geq 1$ always holds, $n^C = 1$ and $P_I(R_4) = \frac{\zeta_c}{\zeta_c+1}$.

Next, recall that as $R \to 0^+$, $n^D \to \infty$, implying that $p(n^D, 0) = \frac{\zeta_d}{\zeta_d+n^D} \to 0$. Since the defection equilibrium exists for all $R > 0$ sufficiently close to 0, it must hold that $P_I(R) \leq p(n^D, 0)$ for such R, and thus that $P_I(R) \to 0$ as $R \to 0^+$. Hence, there exists some R_3 sufficiently close to 0 such that $P_I(R_3) < P_I(R_4)$.

Proof of Lemma 9

Since $P_I(R)$ is non-monotone, there exist $R_1 < R_2$ such that $P_I(R_1) > P_I(R_2)$. Setting $R^{max} = R_2$, we have $R_1 \in (0, R^{max}]$ and $P_I(R_1) > P_I(R^{max})$, so $R^{max} \notin \arg\max_{R \in (0, R^{max}]} P_I(R)$. $\square$

Lemma A.31. ***Within-equilibrium repression effects***

A.31.1 $p(n^D, 0)$ *is an increasing function of* R *on the interval* $R \in (0, \bar{R}^D)$.

A.31.2 There exists $\check{R} \in (\bar{R}^C, \frac{\Pi_1}{\zeta_c+1})$ *such that*

(a) $p(n^C, n^C)$ *is a decreasing function of* R *on the interval* $R \in (\bar{R}^C, \check{R})$ *and*

(b) $p(n^C, n^C)$ *is an increasing function of* R *on the interval* $R \in (\check{R}, \frac{\Pi_1}{\zeta_c+1})$.

Proof:

A.31.1: Let $n^D(R)$ be a function of R, and consider any $R_1, R_2 \in (0, \bar{R}^D)$. From Lemma 2 it follows that $R_1 < R_2 \Rightarrow n^D(R_1) \geq n^D(R_2)$, which in turn implies $p(n^D(R_1), 0) \geq p(n^D(R_2), 0)$.

A.31.2: Let $n^C(R)$ be a function of R, and let $\check{R}$ be as in Lemma 8.

For (a), consider any $R_1, R_2 \in (\check{R}, \frac{\Pi_1}{\zeta_c+1})$. The incumbent's probability of winning the contest in the cooperation equilibrium is $p(n^C(R), n^C(R))$. From Lemma 8 it then follows that

$$\begin{aligned}\check{R} < R_1 < R_2 &\Rightarrow n^C(R_1) \geq n^C(R_2) \\ &\Rightarrow p(n^C(R_1), n^C(R_1)) \leq p(n^C(R_2), n^C(R_2)),\end{aligned}$$

and, symmetrically,

$$\begin{aligned}\check{R} < R_2 < R_1 &\Rightarrow n^C(R_2) \geq n^C(R_1) \\ &\Rightarrow p(n^C(R_1), n^C(R_1)) \geq p(n^C(R_2), n^C(R_2)).\end{aligned}$$

For (b), the argument is analogous, with the final implications

$$\begin{aligned}R_1 < R_2 < \check{R} &\Rightarrow n^C(R_1) \leq n^C(R_2) \\ &\Rightarrow p(n^C(R_1), n^C(R_1)) \geq p(n^C(R_2), n^C(R_2)),\end{aligned}$$

and, symmetrically,

$$\begin{aligned} R_2 < R_1 < \check{R} &\Rightarrow n^C(R_2) \leq n^C(R_1) \\ &\Rightarrow p(n^C(R_1), n^C(R_1)) \leq p(n^C(R_2), n^C(R_2)). \end{aligned}$$

$\square$

Lemma A.32. ***Defection dominates cooperation at $\bar{R}^D$ for a sufficiently large advantage ratio***

Consider $R = \bar{R}^D$ and any $\Pi_1 > \Pi_0 \geq 0$ and $\beta \in (0,1)$. If ζ_d/ζ_c is sufficiently large, then $p(n^D, 0) > p(n^C, n^C)$.

Proof: First, note that at $R = \bar{R}^D$, $n^D \leq 3$. This can be seen by noting that Lemma A.10 implies that $R > \bar{R}^D \Rightarrow n^D \leq 2$, which is equivalent to $R > \bar{R}^D \Rightarrow \tilde{n}^D < 2$. Since $\tilde{n}^D$ is continuous, $R = \bar{R}^D \Rightarrow \tilde{n}^D \leq 2 \Rightarrow n^D \leq 3$.

Next, $\bar{R}^D < \frac{\Pi_1}{\zeta_d+1} < \frac{\Pi_1}{\zeta_c+1}$ implies $\tilde{n}_a^C > 0$ and thus $n_a^C \geq 1$. In combination with $n_c^C \geq 1$ for all $R > 0$, it follows that $n^C \geq 1$ at $\bar{R}^D$. Then

$$p(n^D, 0) > p(n^C, n^C) \Leftrightarrow \frac{\zeta_d}{\zeta_d + n^D} > \frac{\zeta_c}{\zeta_c + n^C} \Leftrightarrow \frac{\zeta_d}{\zeta_c} > \frac{n^D}{n^C}.$$

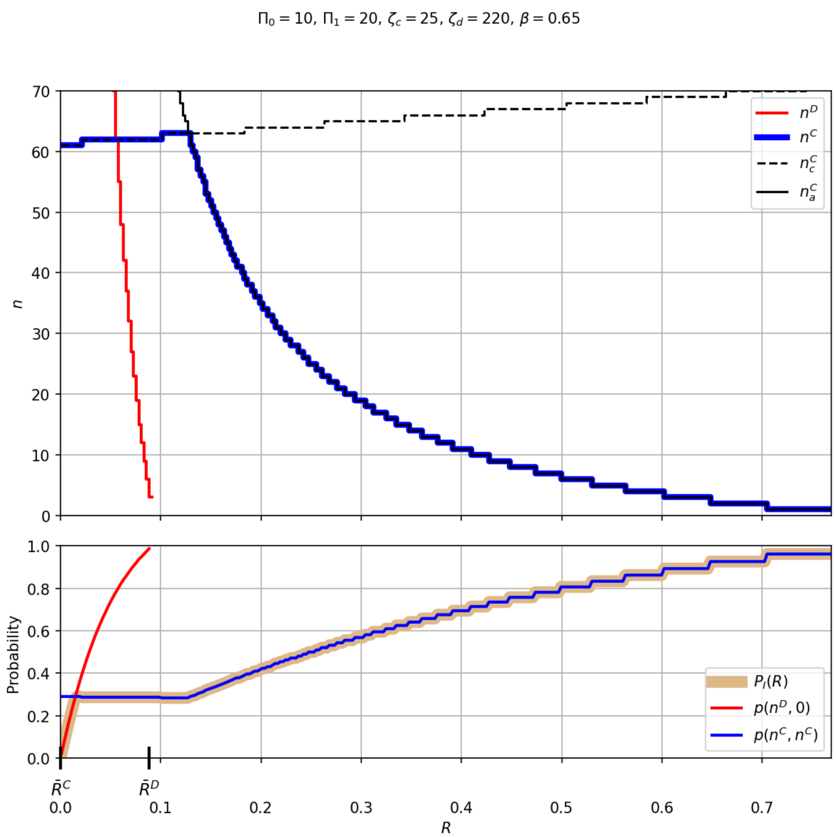
Since n^D/n^C is bounded above by 3, it follows that $\zeta_d/\zeta_c > 3$ implies $p(n^D, 0) > p(n^C, n^C)$ at $\bar{R}^D$, which proves the lemma. $\square$

B Additional Figures

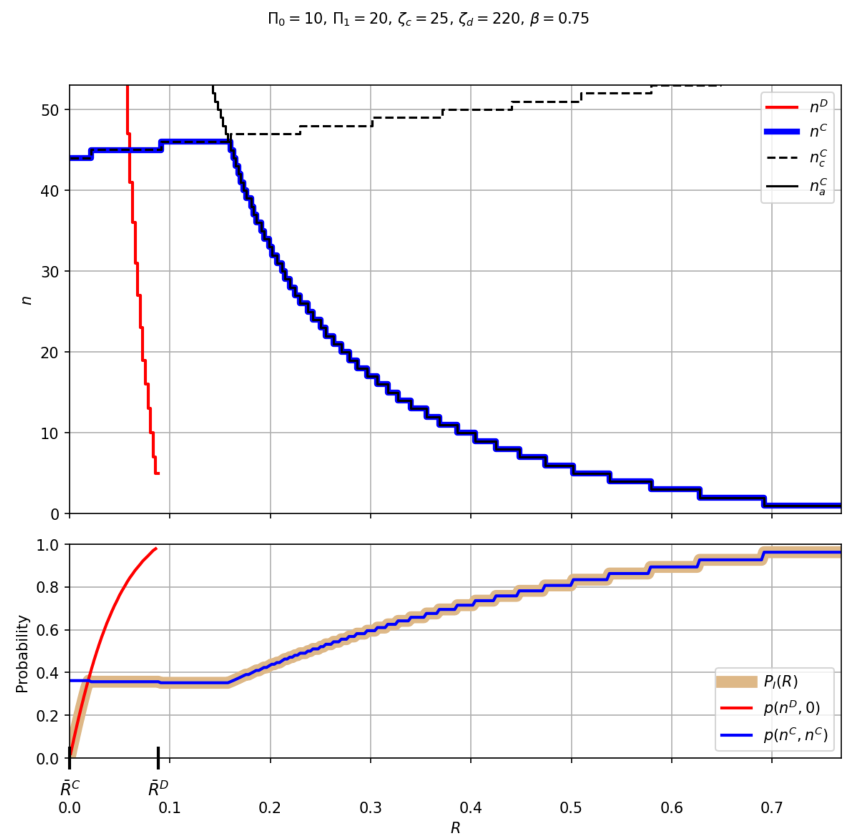
This appendix presents additional simulations of our model to show that the non-monotonicity of $P_I(R)$ exists across a range of parameter values, as well as present examples for which it does not arise. The parameter space spans

- $(\Pi_0, \Pi_1) \in \{(10, 20), (10, 100)\}$,
- $(\zeta_d, \zeta_c) \in \{(220, 25), (200, 45), (180, 65), (160, 105)\}$, and
- $\beta \in \{0.65, 0.75, 0.85, 0.95\}$.

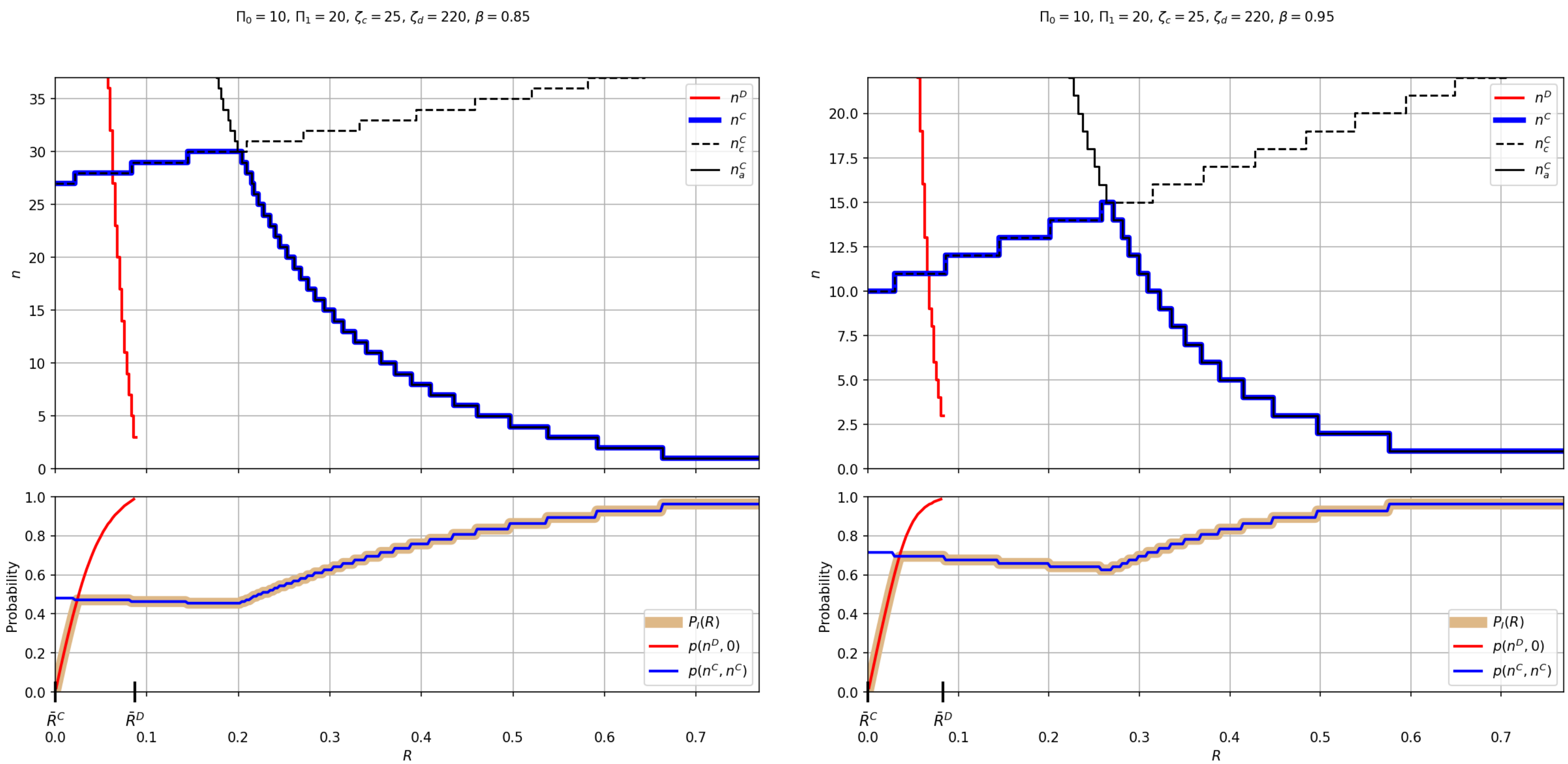
These simulations are illustrative rather than a structural calibration of the model. Source code is available upon request.



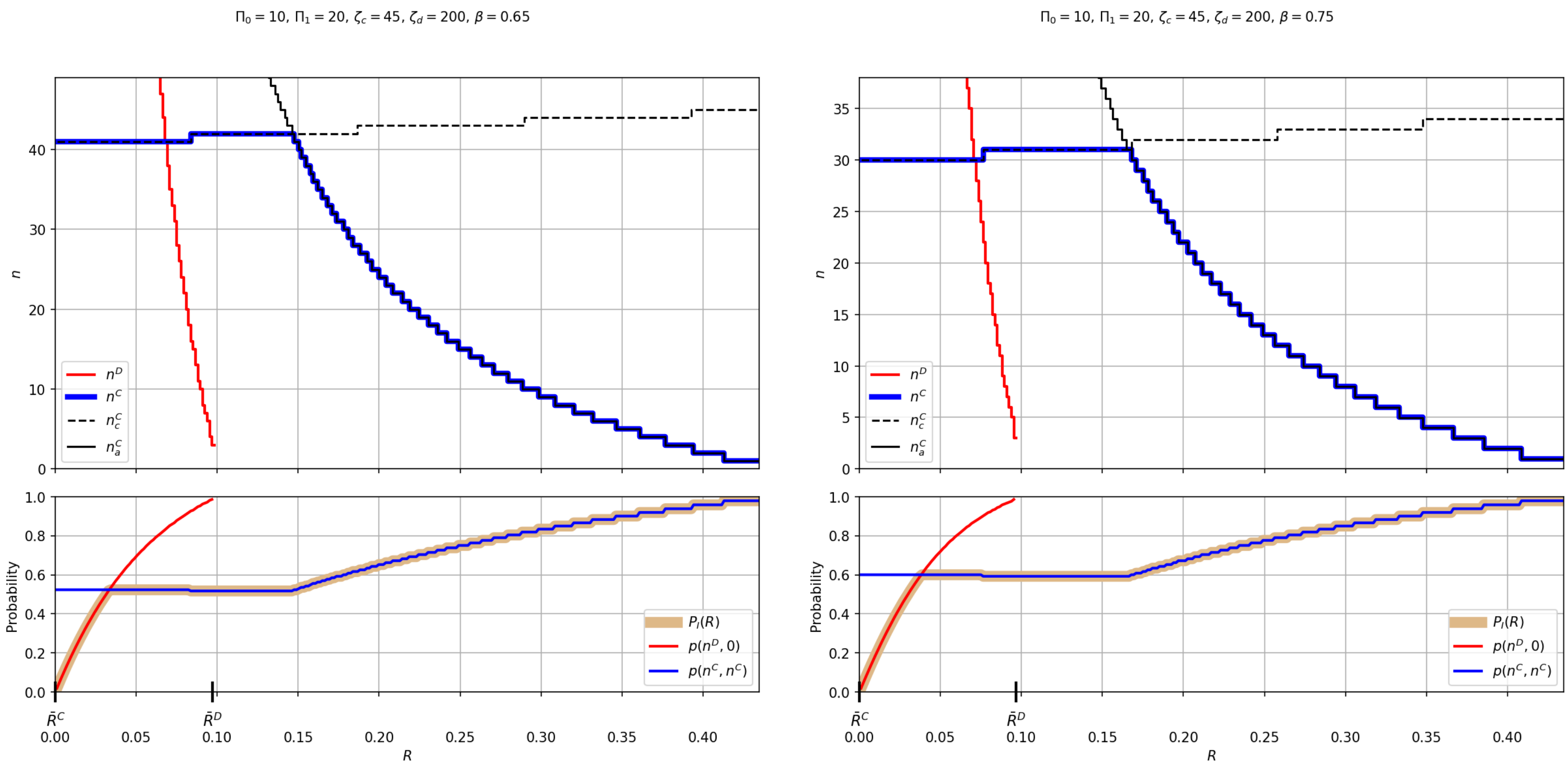
$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 25$, $\zeta_d = 220$, $\beta = 0.65$
n
n^D
n^C
n_c^C
n_a^C
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R



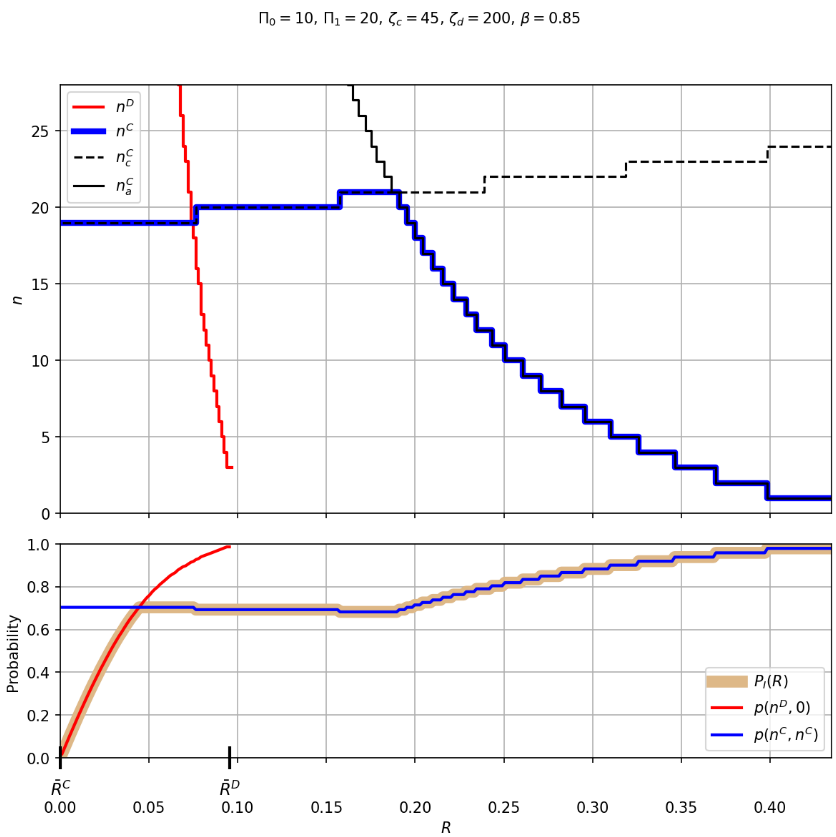
$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 25$, $\zeta_d = 220$, $\beta = 0.75$
n
n^D
n^C
n_c^C
n_a^C
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R



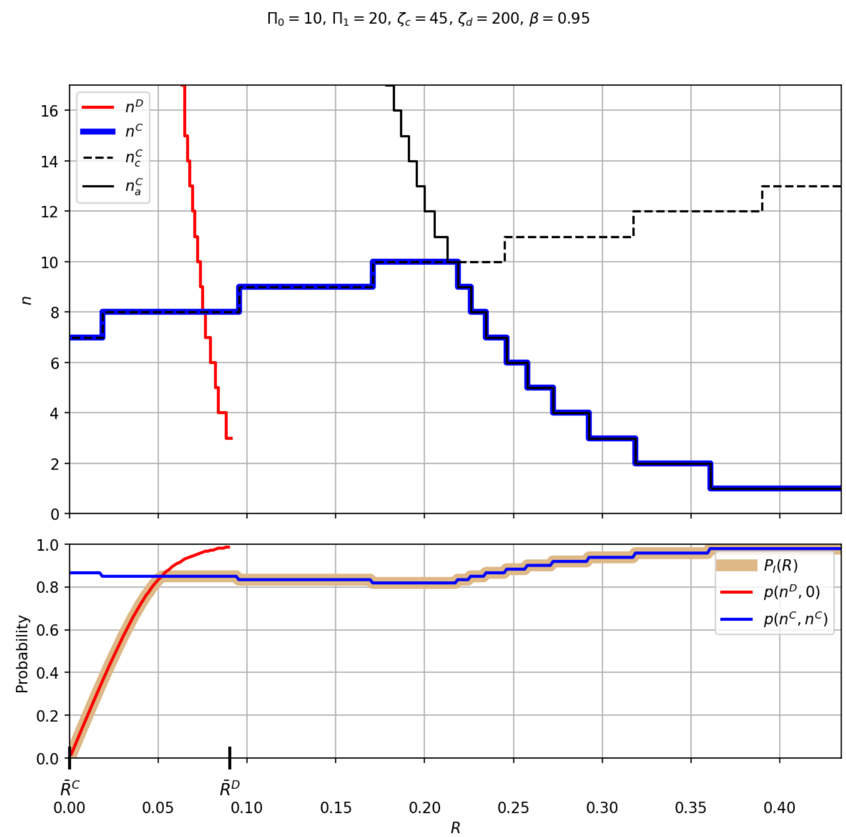
$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 25$, $\zeta_d = 220$, $\beta = 0.85$
$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 25$, $\zeta_d = 220$, $\beta = 0.95$
n^D
n^C
n^C_c
n^C_a
n
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R



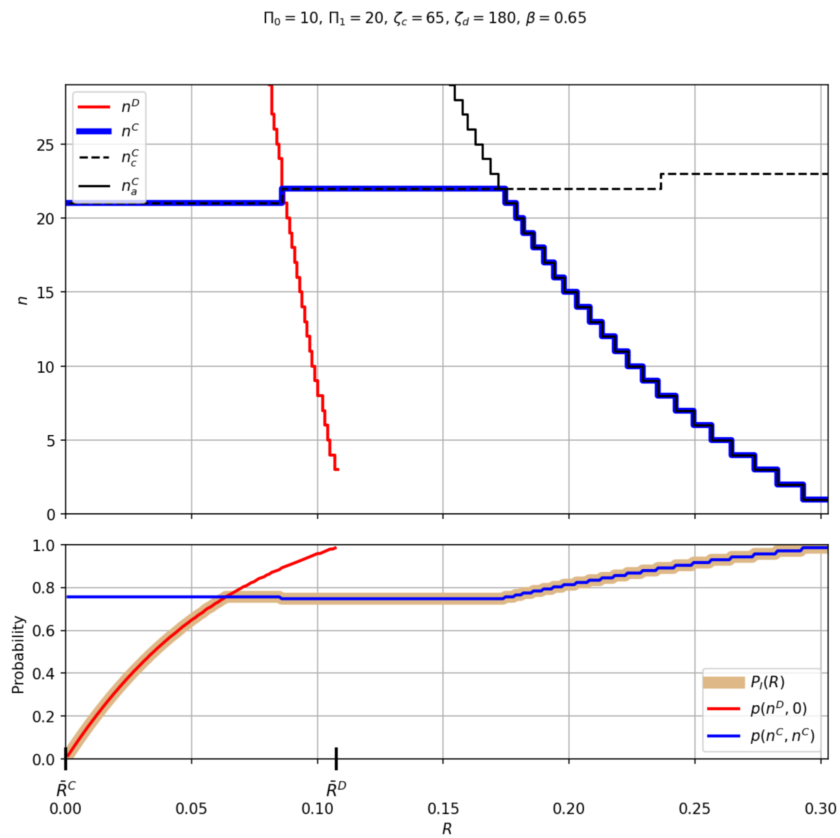
$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 45$, $\zeta_d = 200$, $\beta = 0.65$
$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 45$, $\zeta_d = 200$, $\beta = 0.75$
n
n^D
n^C
n^C_c
n^C_a
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R

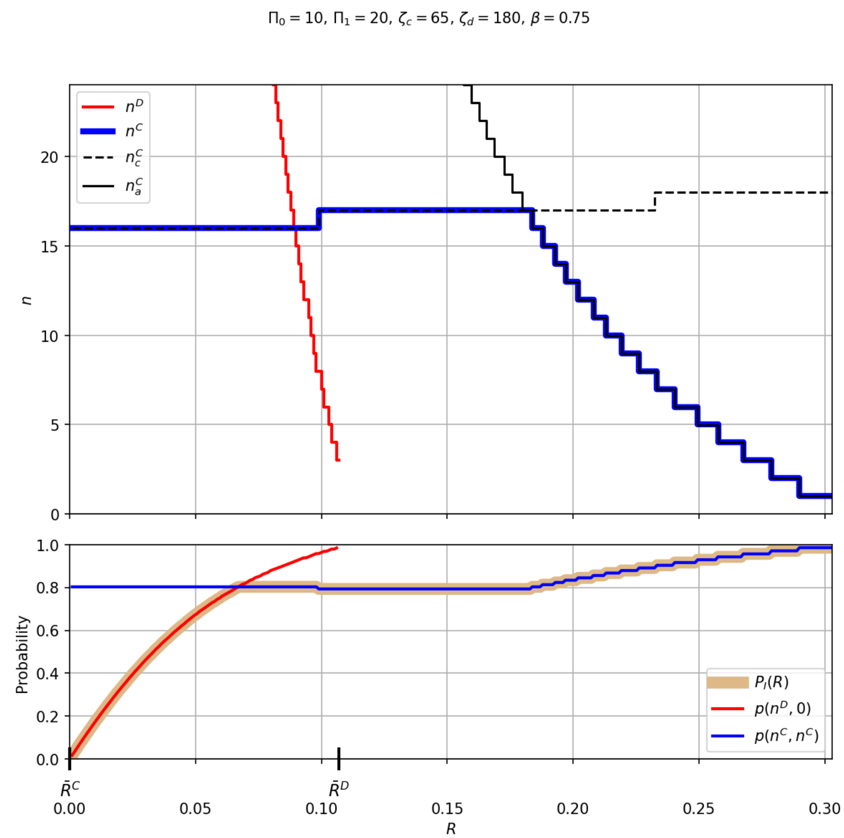


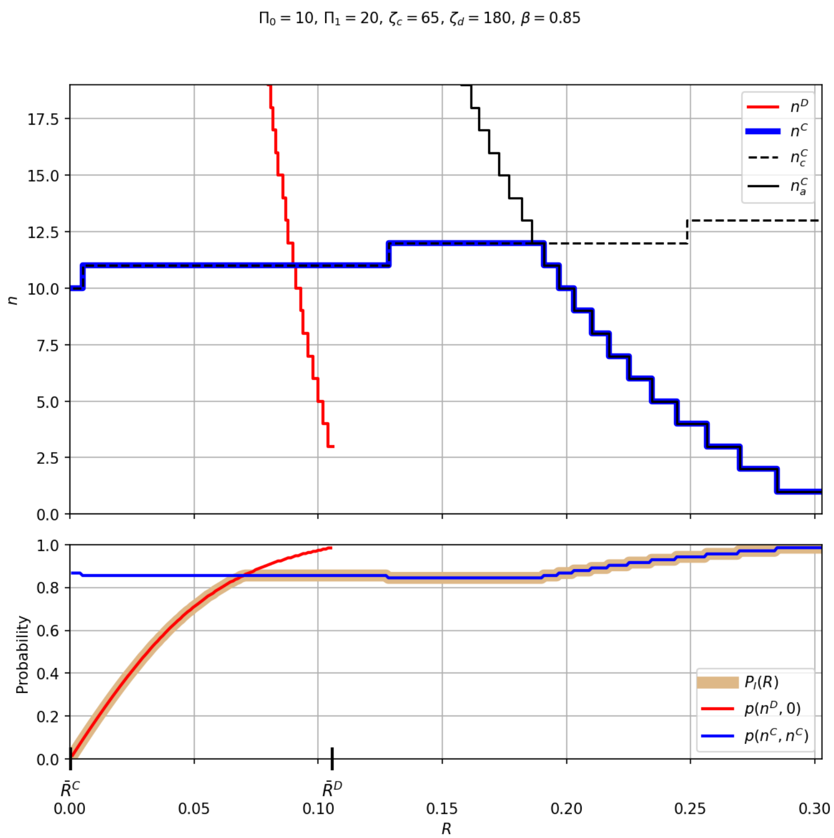
$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 45$, $\zeta_d = 200$, $\beta = 0.85$
n^D
n^C
n_c^C
n_a^C
n
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R

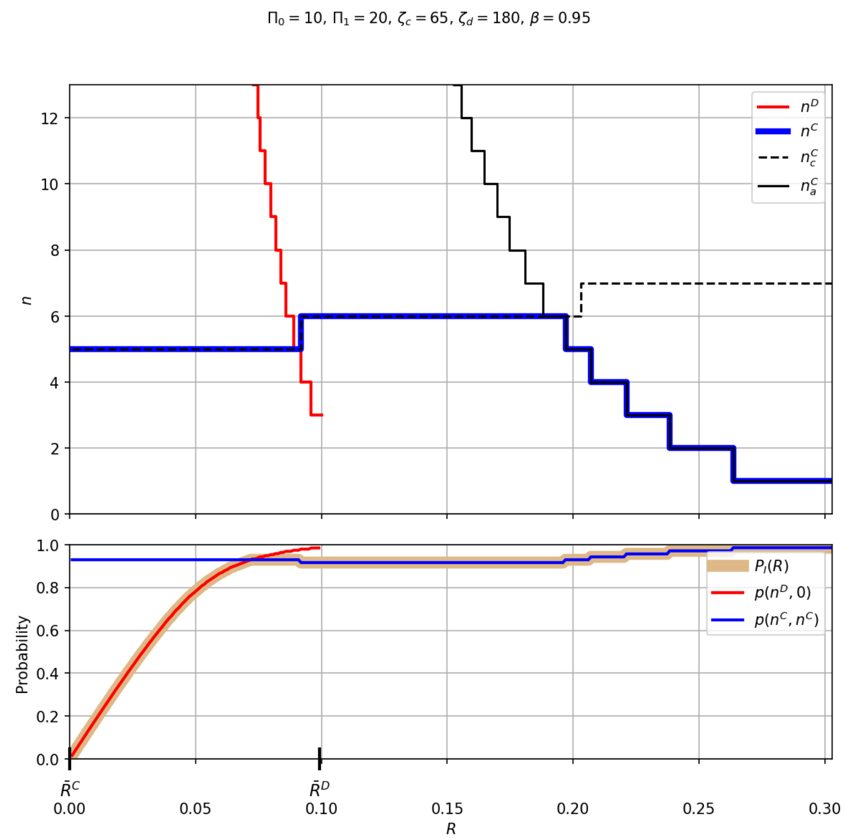


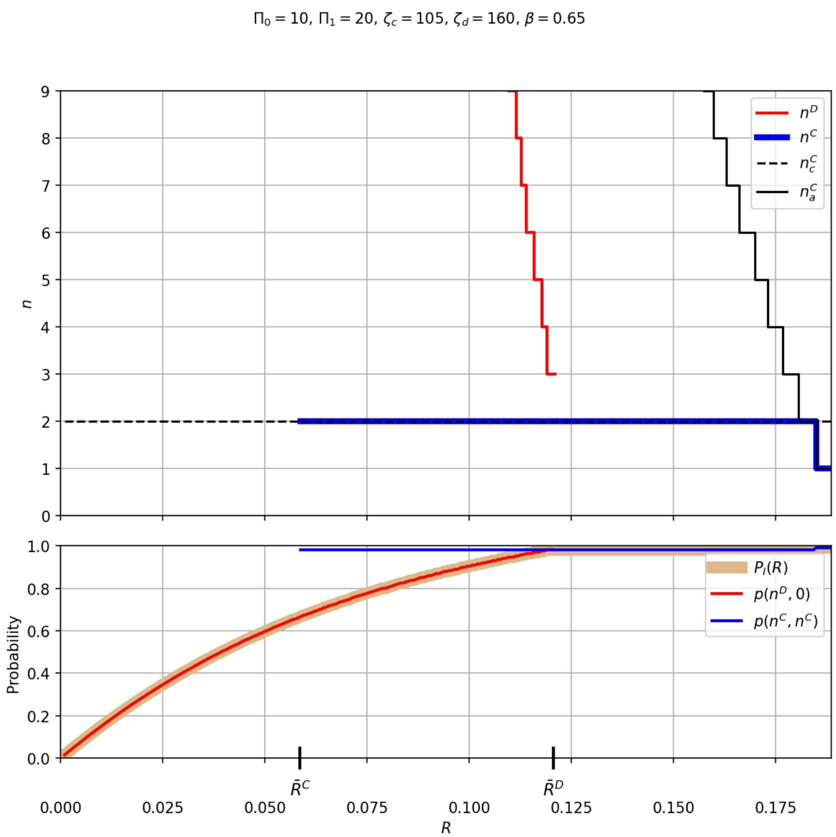
$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 45$, $\zeta_d = 200$, $\beta = 0.95$
n^D
n^C
n_c^C
n_a^C
n
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R

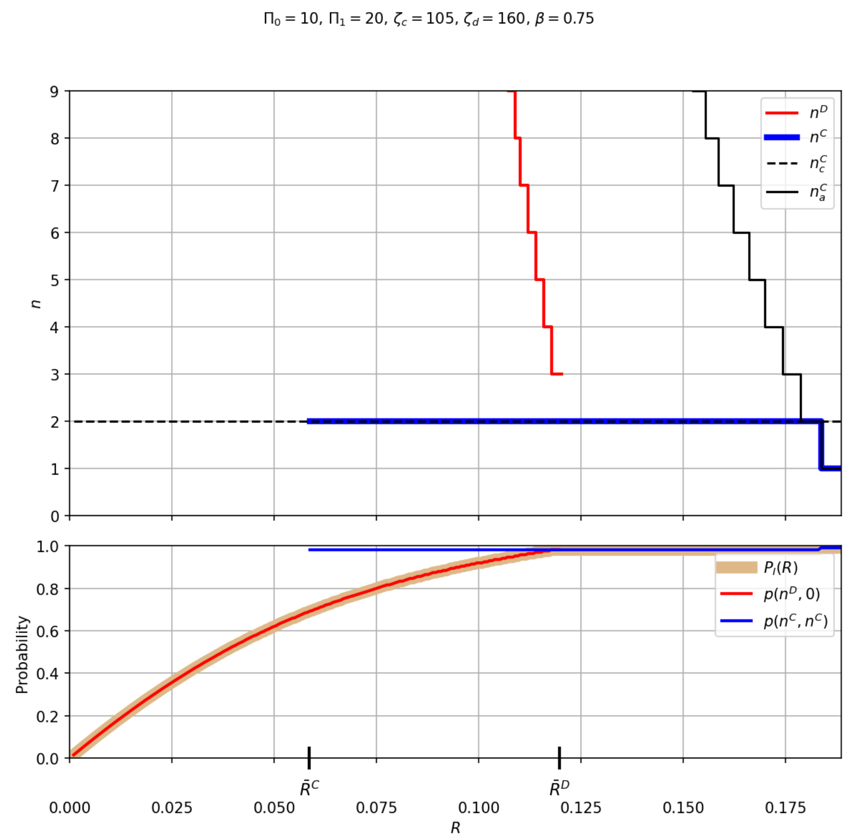

$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 65$, $\zeta_d = 180$, $\beta = 0.65$
n^D
n^C
n^C_c
n^C_a
n
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R

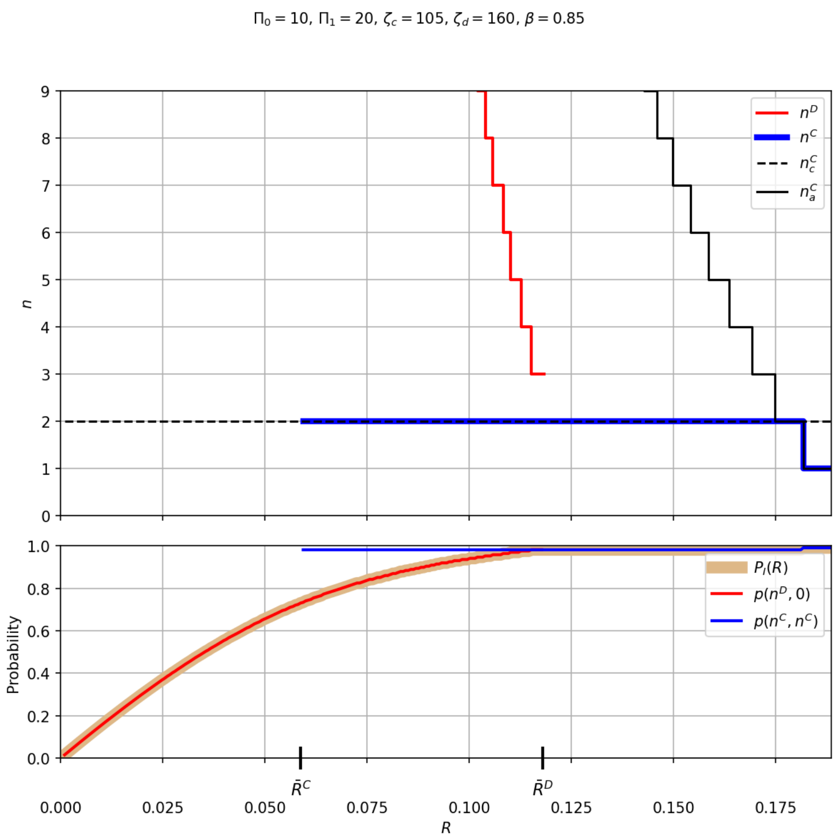

$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 65$, $\zeta_d = 180$, $\beta = 0.75$
n^D
n^C
n^C_c
n^C_a
n
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R

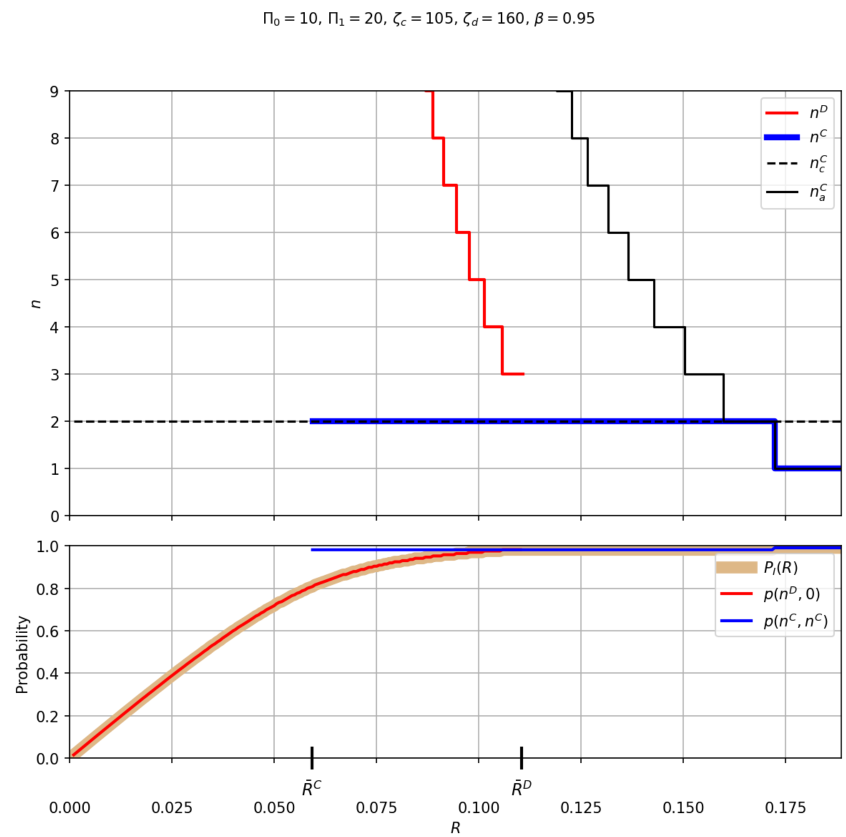

$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 65$, $\zeta_d = 180$, $\beta = 0.85$
n
n^D
n^C
n_c^C
n_a^C
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R

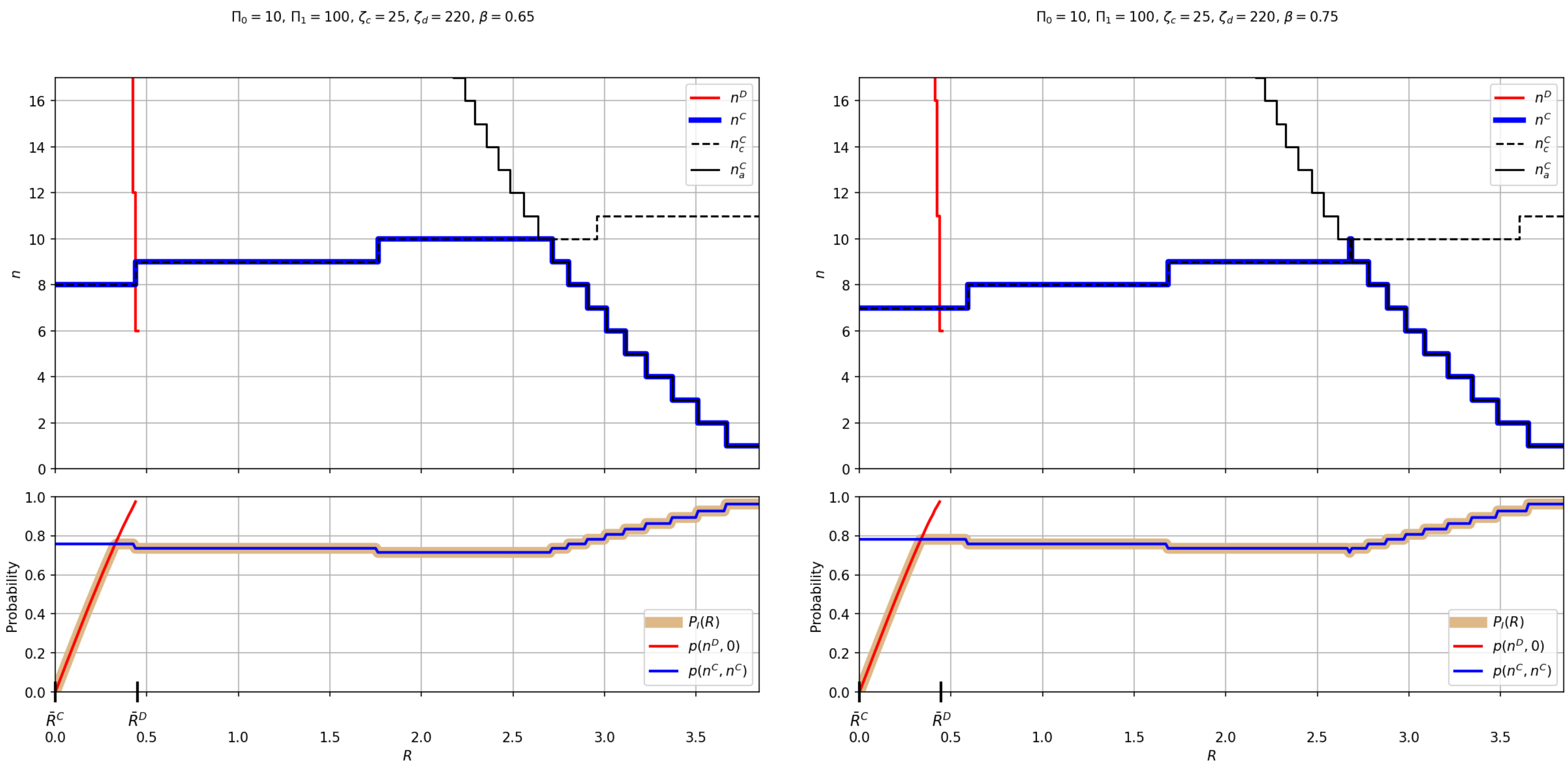

$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 65$, $\zeta_d = 180$, $\beta = 0.95$
n
n^D
n^C
n_c^C
n_a^C
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R


$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 105$, $\zeta_d = 160$, $\beta = 0.65$
n^D
n^C
n_c^C
n_a^C
n
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$
$\bar{R}^D$
R

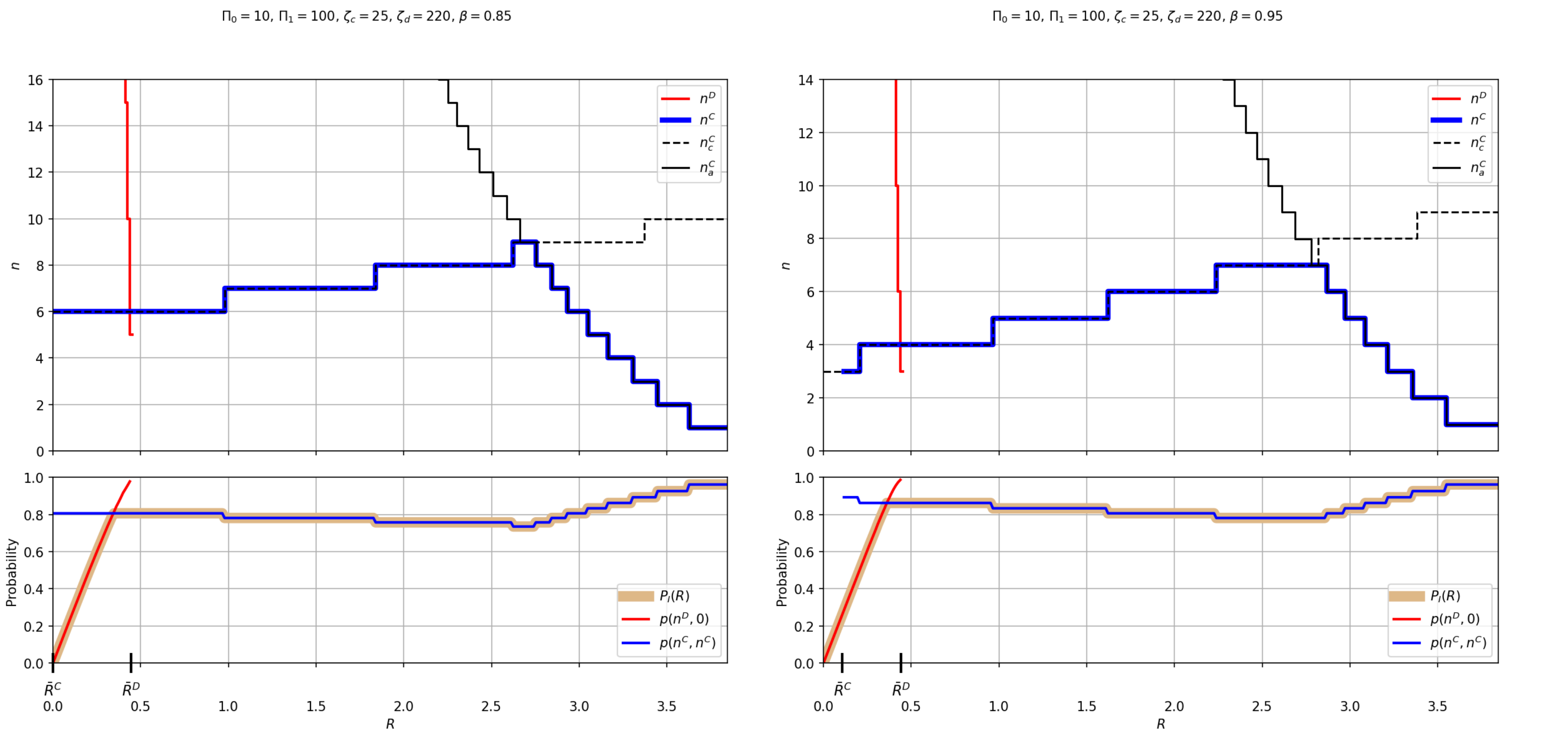

$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 105$, $\zeta_d = 160$, $\beta = 0.75$
n^D
n^C
n_c^C
n_a^C
n
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$
$\bar{R}^D$
R


$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 105$, $\zeta_d = 160$, $\beta = 0.85$
n^D
n^C
n_c^C
n_a^C
n
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$
$\bar{R}^D$
R

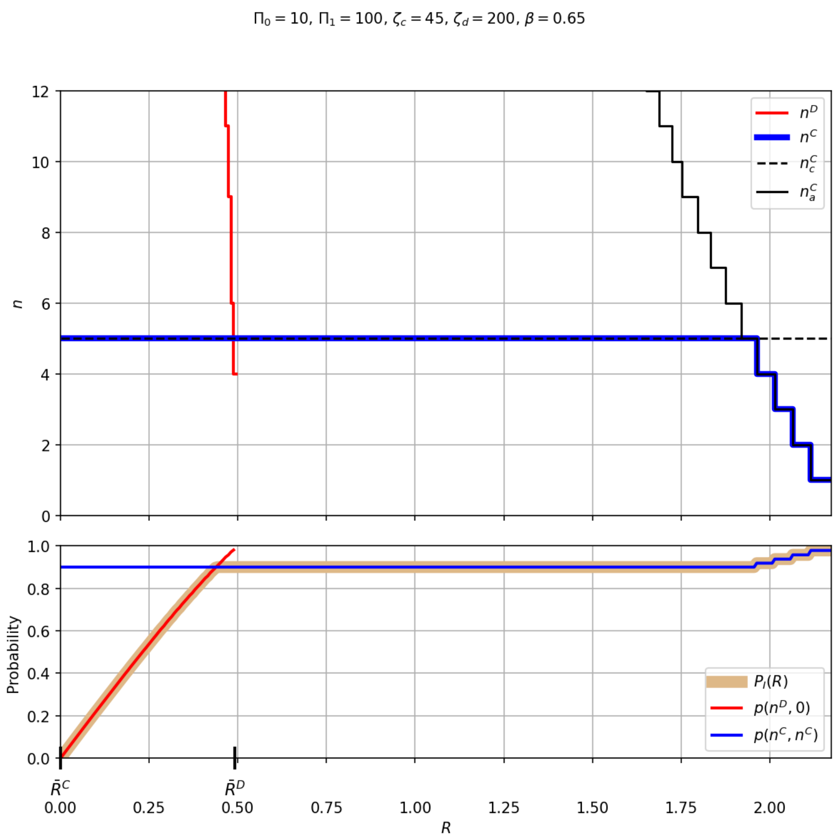

$\Pi_0 = 10$, $\Pi_1 = 20$, $\zeta_c = 105$, $\zeta_d = 160$, $\beta = 0.95$
n^D
n^C
n_c^C
n_a^C
n
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$
$\bar{R}^D$
R

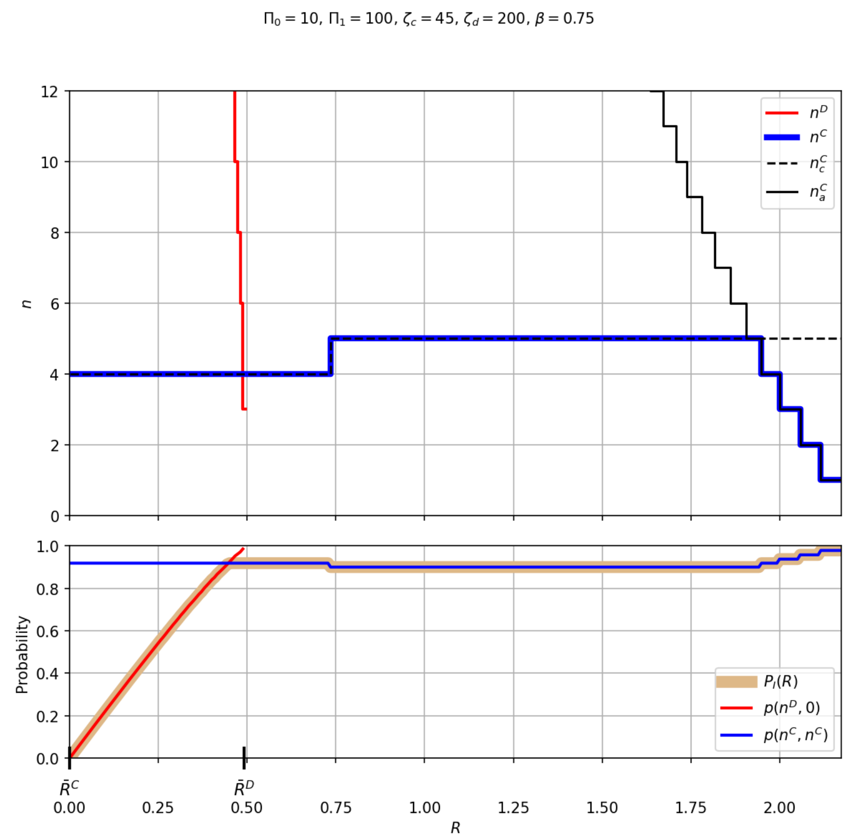


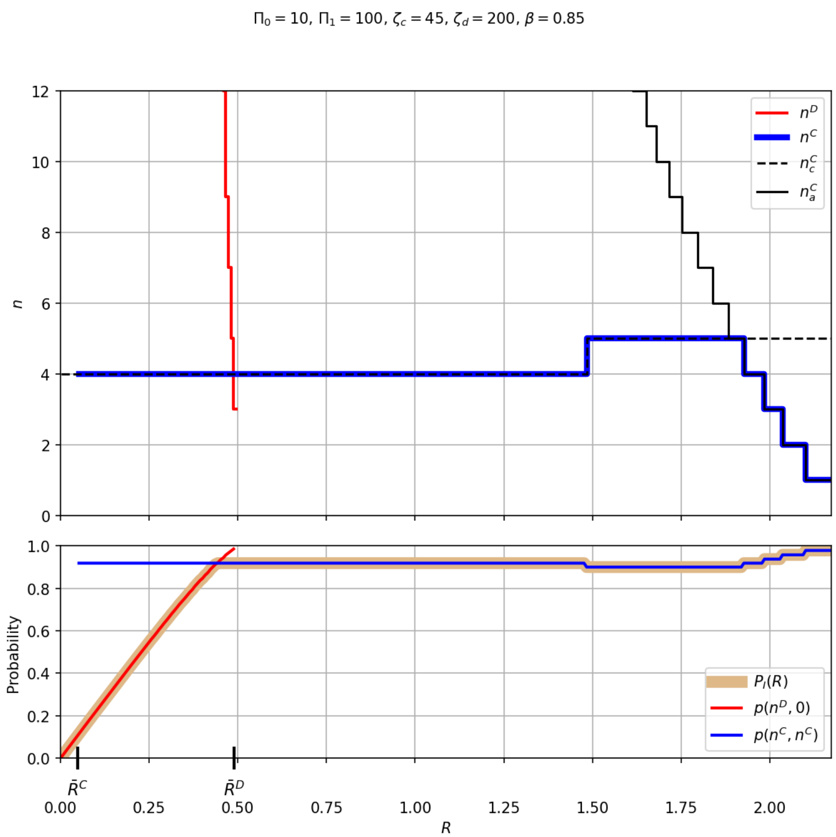
$\Pi_0 = 10,\ \Pi_1 = 100,\ \zeta_c = 25,\ \zeta_d = 220,\ \beta = 0.65$
$\Pi_0 = 10,\ \Pi_1 = 100,\ \zeta_c = 25,\ \zeta_d = 220,\ \beta = 0.75$
n^D
n^C
n^C_c
n^C_a
n
Probability
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R

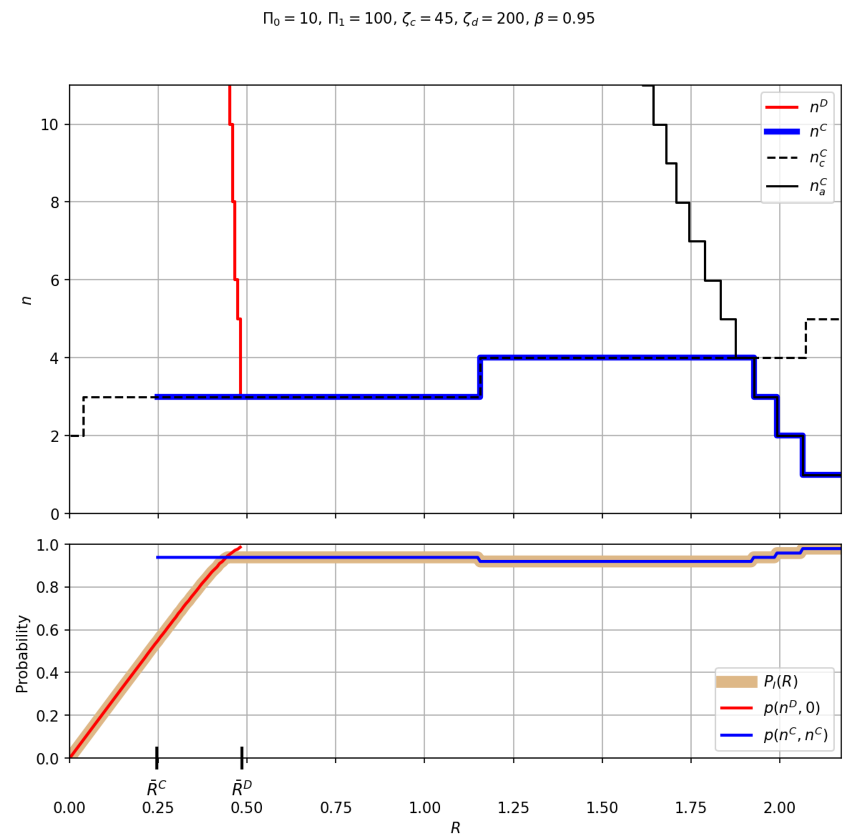


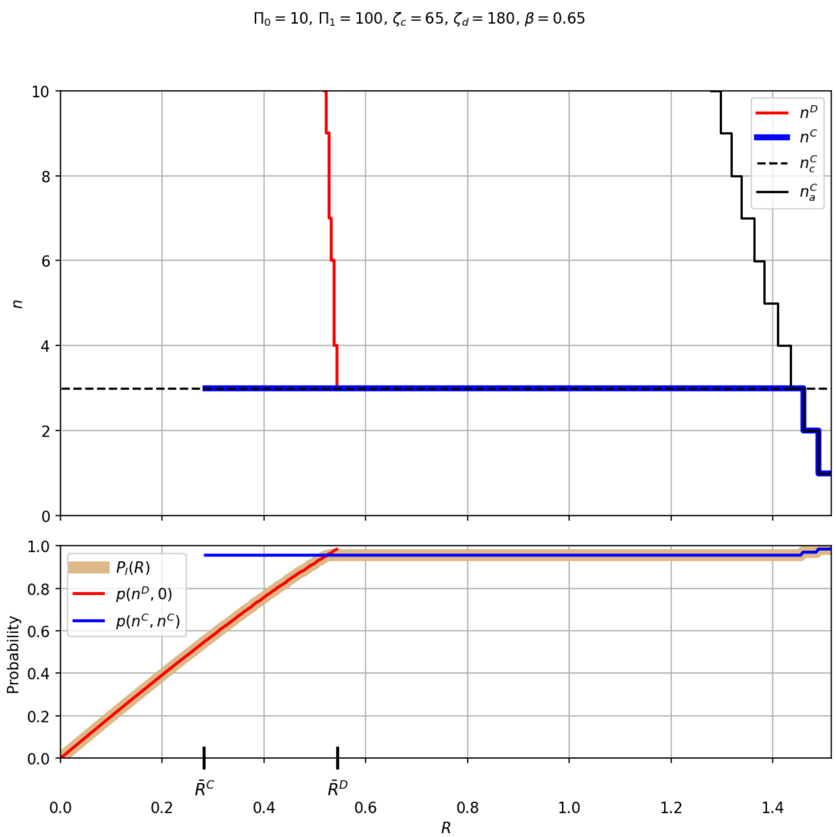
$\Pi_0 = 10,\ \Pi_1 = 100,\ \zeta_c = 25,\ \zeta_d = 220,\ \beta = 0.85$
$\Pi_0 = 10,\ \Pi_1 = 100,\ \zeta_c = 25,\ \zeta_d = 220,\ \beta = 0.95$
n
n^D
n^C
n^C_c
n^C_a
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R

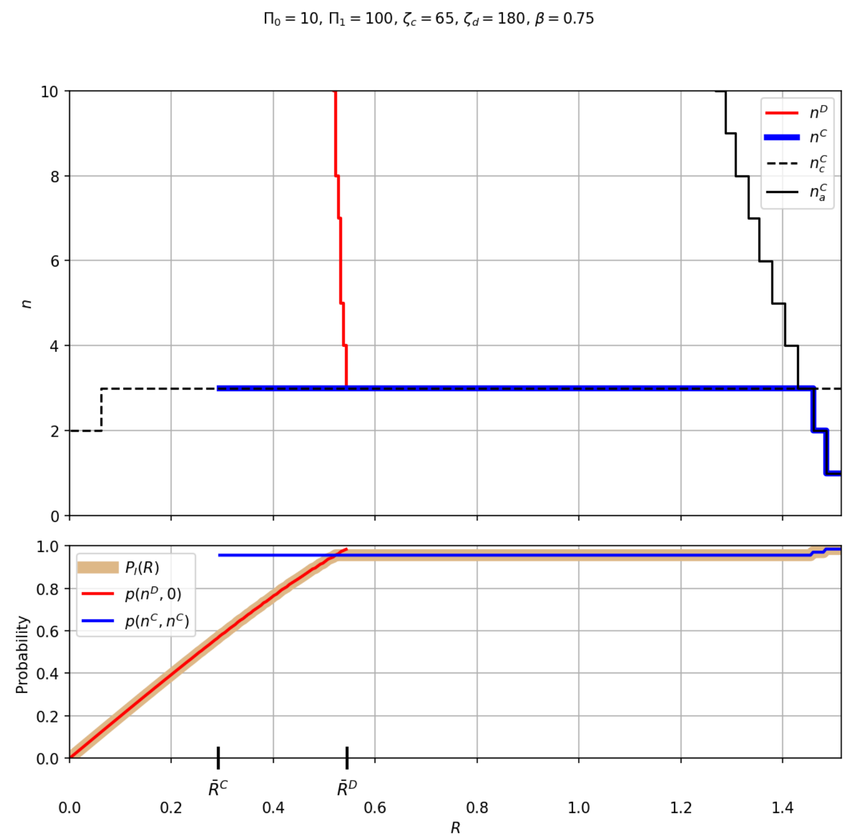

$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 45$, $\zeta_d = 200$, $\beta = 0.65$
n
n^D
n^C
n_c^C
n_a^C
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R

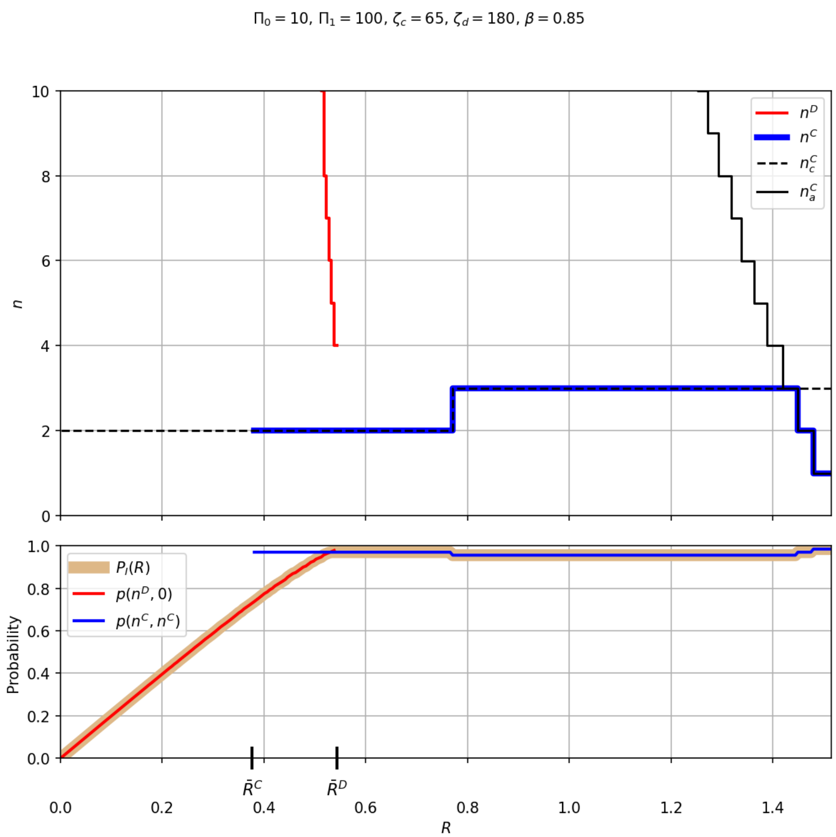

$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 45$, $\zeta_d = 200$, $\beta = 0.75$
n
n^D
n^C
n_c^C
n_a^C
Probability
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
$\bar{R}^C$
$\bar{R}^D$
R


$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 45$, $\zeta_d = 200$, $\beta = 0.85$
n^D
n^C
n_c^C
n_a^C
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
n
R
$\bar{R}^C$
$\bar{R}^D$

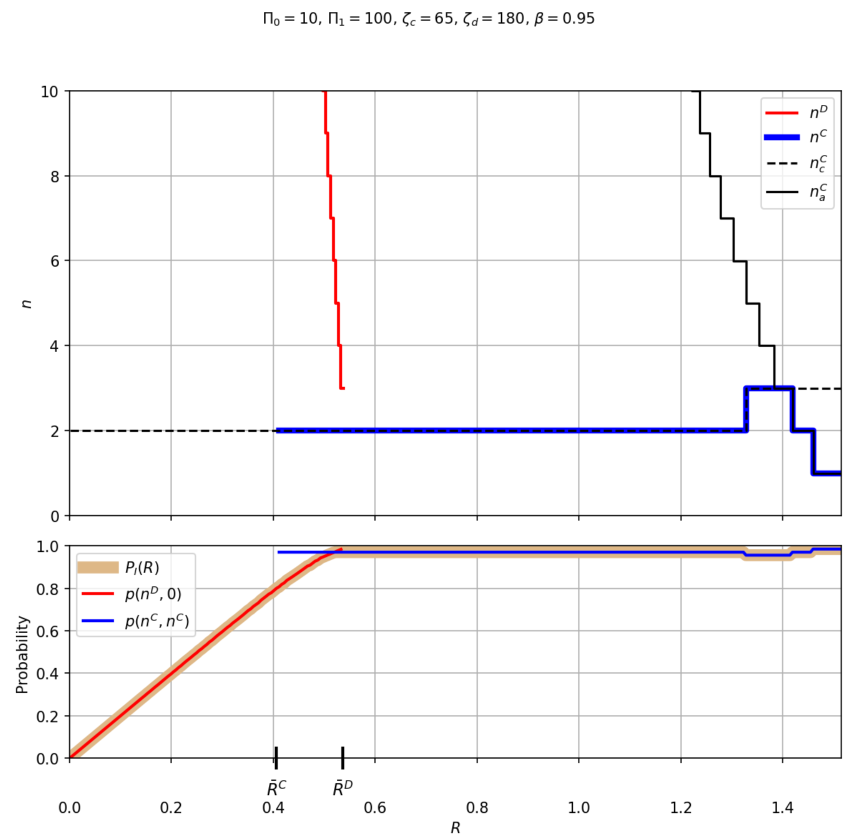

$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 45$, $\zeta_d = 200$, $\beta = 0.95$
n^D
n^C
n_c^C
n_a^C
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
n
R
$\bar{R}^C$
$\bar{R}^D$


$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 65$, $\zeta_d = 180$, $\beta = 0.65$
n^D
n^C
n^C_c
n^C_a
n
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$
$\bar{R}^D$
R

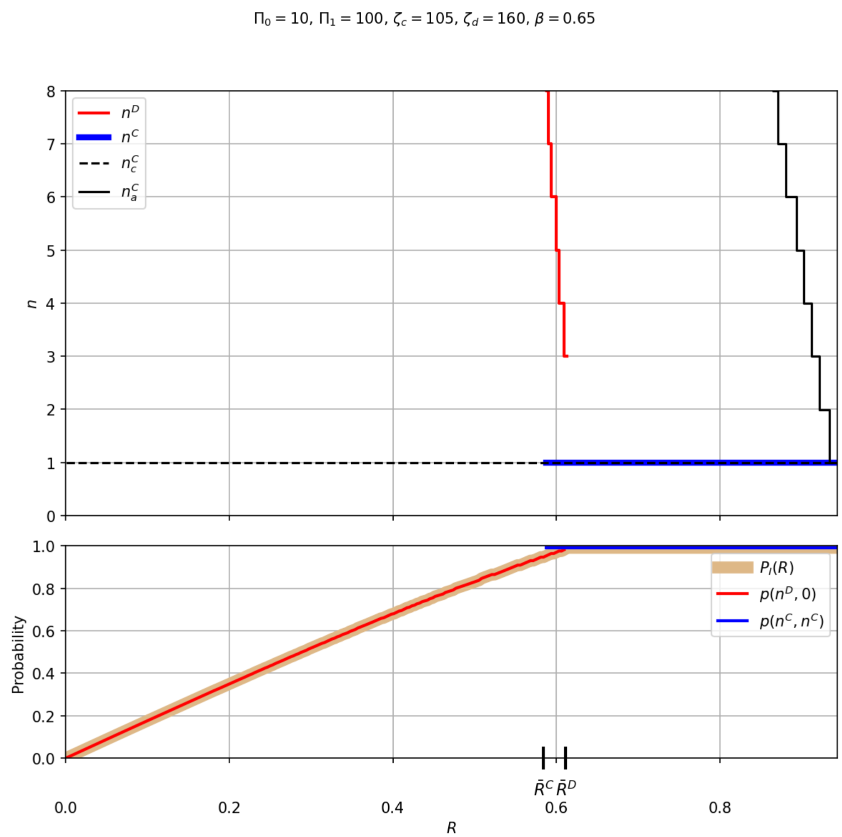

$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 65$, $\zeta_d = 180$, $\beta = 0.75$
n^D
n^C
n^C_c
n^C_a
n
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$
$\bar{R}^D$
R

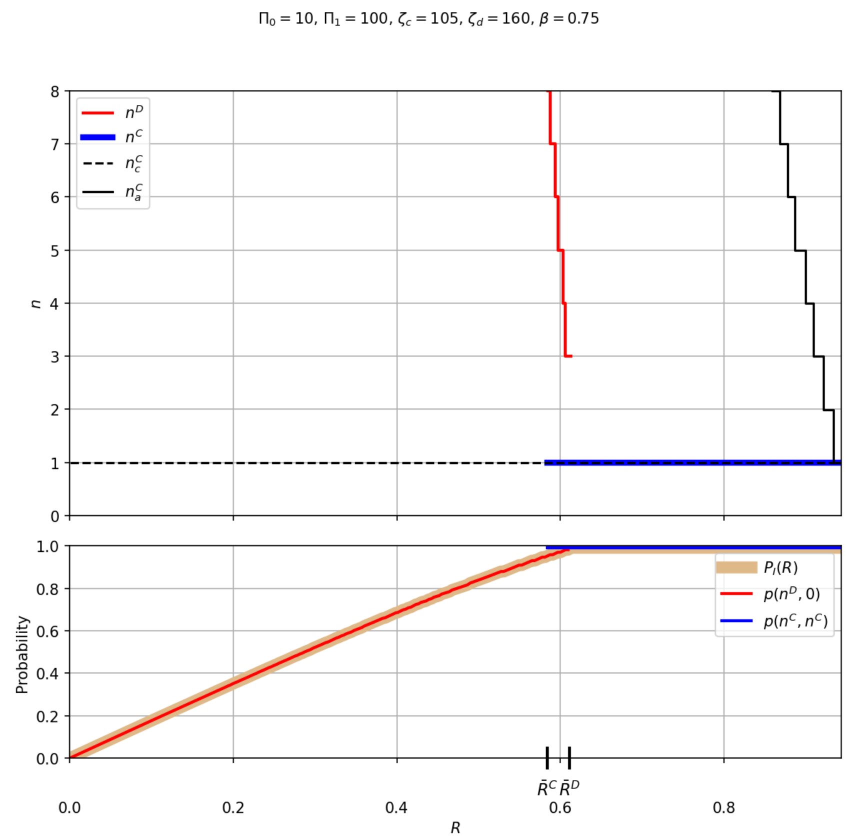


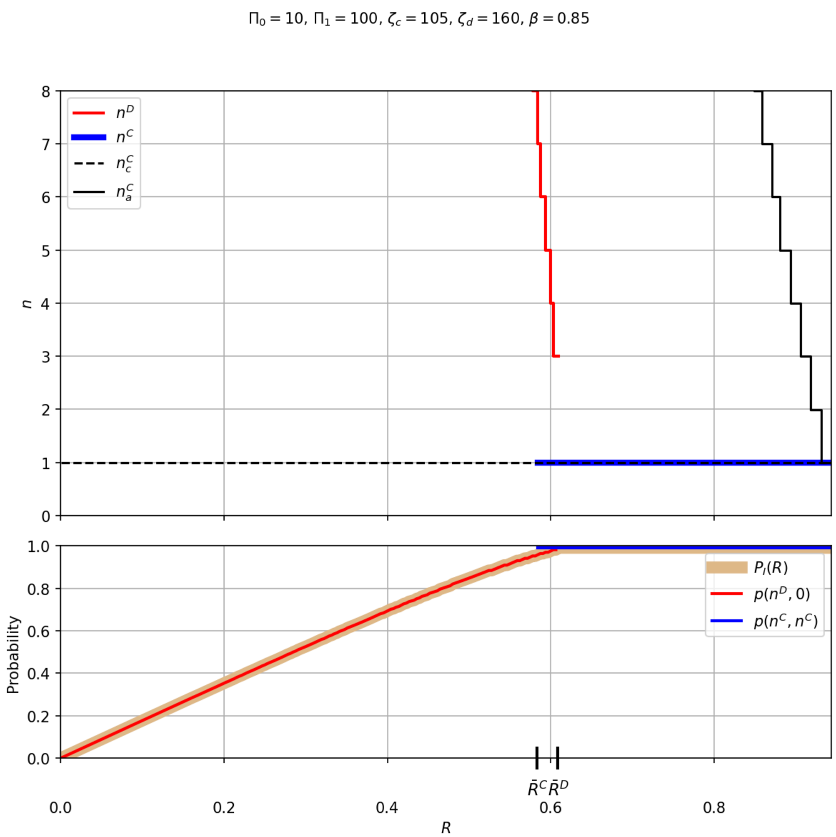
$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 65$, $\zeta_d = 180$, $\beta = 0.85$
n^D
n^C
n^C_c
n^C_a
n
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$
$\bar{R}^D$
R

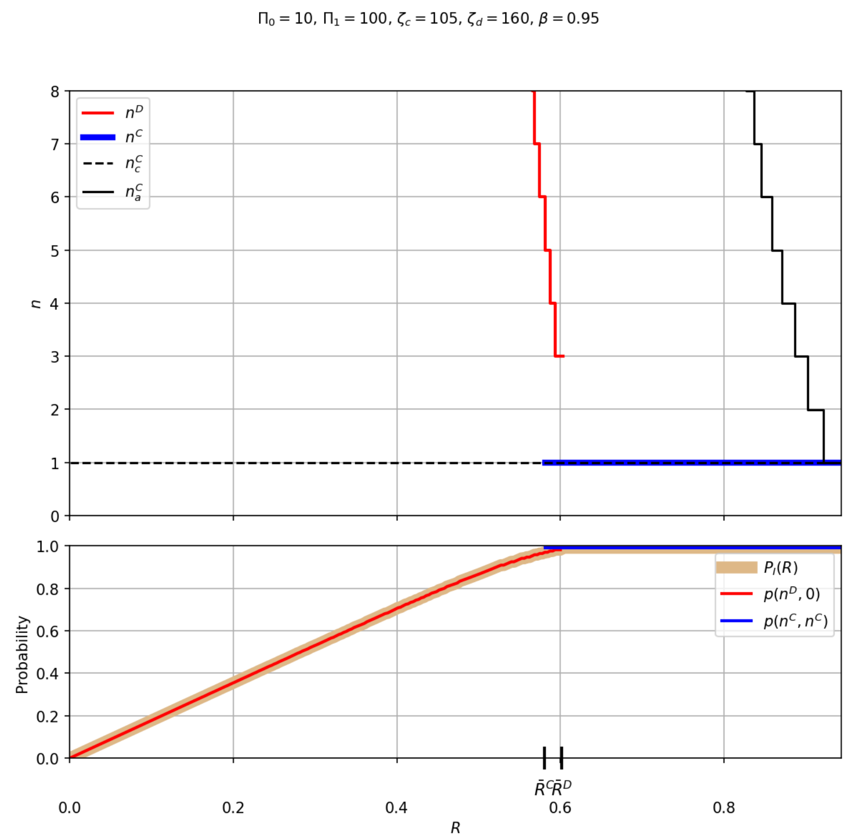


$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 65$, $\zeta_d = 180$, $\beta = 0.95$
n^D
n^C
n^C_c
n^C_a
n
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$
$\bar{R}^D$
R


$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 105$, $\zeta_d = 160$, $\beta = 0.65$
n^D
n^C
n_c^C
n_a^C
n
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$ $\bar{R}^D$
R


$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 105$, $\zeta_d = 160$, $\beta = 0.75$
n^D
n^C
n_c^C
n_a^C
n
$P_l(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$ $\bar{R}^D$
R


$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 105$, $\zeta_d = 160$, $\beta = 0.85$
n^D
n^C
n_c^C
n_a^C
n
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$ $\bar{R}^D$
R


$\Pi_0 = 10$, $\Pi_1 = 100$, $\zeta_c = 105$, $\zeta_d = 160$, $\beta = 0.95$
n^D
n^C
n_c^C
n_a^C
n
$P_I(R)$
$p(n^D, 0)$
$p(n^C, n^C)$
Probability
$\bar{R}^C$ $\bar{R}^D$
R