Equilibrium Trade Regimes: Power- vs. Rules-Based*

Cecilia Carvalho† Daniel Monte‡ Emanuel Ornelas§

November 18, 2025

Abstract

We study the sustainability of a rules-based trade regime in a dynamic framework where the leading country sets the trade regime and leadership changes over time. Transitioning from power to rules requires a hegemon. Once established, rules may be upheld by non-hegemonic leaders anticipating future loss of dominance. A rules-based equilibrium is viable only when costs of establishing it are moderate, countries are sufficiently patient, and leadership turnover is high. We highlight the trade-offs that a redesign of the WTO must confront to remain viable.

Keywords: Hegemonic Stability Theory; World Trade Organization; Trade Agreements

JEL classification: F02; F13; F53

*We are grateful to Kyle Bagwell, Stephen Brooks, Peter Karrel, Nuno Limao, Petros Mavroidis, Stratos Pahis, Tom Prusa and Gregory Shaffer for detailed discussions of our paper. We also thank Paula Bustos, Camila Campos, Andrés Carvajal, Paola Conconi, James Fenske, Sebastian Krautheim, Timothy Meyer, Isabelle Mejean, Joao Paulo Pessoa, Dani Rodrik, David Rubin, Bob Staiger, Ben Zissimos, and seminar participants at the 2025 CESifo Area Conference on Global Economy, 2025 ERWIT-CEPR, 2024 GEP/CEPR Postgraduate Conference, 2024 Gravity, Trade Agreements and Policy Workshop, 2023 InsTED Workshop, 2025 KIEL-CEPR Geoeconomics Workshop, 2024 Rethinking the Global Trading System Workshop, *Staigerfest* Workshop, 2024 TIGN Workshop, 2025 Villars Workshop on International Trade and Trade Policy, 2023 World Congress of the International Economic Organization, Carlos III, EPGE, Geneva Trade and Development Workshop, INSPER, Online Australasian Seminar in International Economics, Paris School of Economics, Sao Paulo School of Economics, USP and U. Tennessee for valuable comments and discussions. We thank Nicolas Moura for exceptional research assistance. Ornelas acknowledges financial support from CNPq.

†University of Geneva (GSEM) and HES-SO/HEG-Genève. E-mail: cecilia.menezes@etu.unige.ch

‡University of Turin, ESOMAS Department, Collegio Carlo Alberto, and Sao Paulo School of Economics-FGV. E-mail: daniel.monte@unito.it.

§Sao Paulo School of Economics-FGV, CEPR, CESifo & CEP-LSE. E-mail: eaornelas@gmail.com.

1 Introduction

Led by the United States, the multilateral trading system has operated under a rules-based regime since the late 1940s, culminating with the creation of the World Trade Organization (WTO) in 1995. Currently, the WTO has 166 member countries, which are responsible for around 98% of world trade. Despite its apparent complexity, until recently it worked remarkably well, with members generally following the established rules and abiding by the decisions of the dispute resolution system. However, in the last 10-15 years, the system has shown alarming signs of weakness, with members increasingly breaking rules and disputes being left unresolved. The recent trade wars and the waning hegemonic power of the U.S. in the global scenario, in parallel with the rise of China, have reinforced the uncertainties about the future of the rules-based system. In particular, the U.S. appears no longer willing to underwrite an international trading system based on rules, as it did for several decades, and China has not shown any desire to take on that role. To understand these changes and their implications, we study the sustainability of a rules-based trade regime when the leading country's relative importance and identity change over time.

We model the choice of the trading system – whether based on rules or on power – as the prerogative of the world's largest economy, whose identity evolves stochastically. We show that the viability of a lasting rules-based regime depends on a delicate set of conditions. When the cost of establishing such a system is low, the world enters a cyclical equilibrium, alternating between rules-based and power-based regimes. When that cost is high, a power-based equilibrium persists, even if collectively countries would benefit from a rules-based regime. Only when the cost is moderate and countries are sufficiently patient (i.e., have high discount factors) can a rules-based regime be sustained in equilibrium. But its sustainability also requires sufficient turnover in global leadership, so that dominant countries maintain the rules-based system in anticipation of losing leadership – and thereby avoid future exploitation under a power-based regime. Furthermore, if the world reaches a state in which there are two or more concurrent leaders, other difficulties arise. In particular, free-riding and market-power incentives tend to make the sustainability of a rules-based equilibrium even harder.

To focus on the long-run viability of trading regimes, we make assumptions about which regime the "leader" country prefers in the short run, following the existing literature to link these preferences to country size. Large, "dominant" countries prefer to operate under a power-based system, because it allows them to extract greater gains from trade by leveraging their size in negotiations. This presumes that smaller countries are willing to participate. However, as size asymmetries increase – and the leading country becomes a "hegemon" – smaller countries may come to expect ex-post exploitation to such an extent that they prefer to abstain from power-based arrangements. In that case, a rules-based regime becomes the best alternative for the hegemon, as it ensures the participation of smaller countries.[1]

Those static regime preferences align with the messages from the *Hegemonic Stability Theory* (HST), initially proposed by Kindleberger (1973). According to the HST, there is a cost to providing an international public good and, given the lack of coordination between countries and the absence of a coercive international system, its provision will generally be inefficiently low. This changes when there is a hegemon – an economy significantly larger than all others. Being able to capture a share of the benefit of the public good greater than the total cost of providing it, the hegemon is willing to bear these costs. In this sense, the presence of a hegemon fosters the establishment and maintenance of a rules-based international regime. And as smaller nations share the benefits from this arrangement but not the costs, this system yields collectively desirable results for all countries. However, it relies on the existence of a hegemon; in the absence of hegemonic power, global disorder and inefficient outcomes prevail. The HST therefore predicts that a rules-based trade regime will endure only while there is a hegemon to sponsor it. Based on these insights, it associates the establishment and sustainability of liberal trading regimes in the mid-19th and mid-20th centuries with the British and American hegemonic leaderships, respectively, and the shift to protectionism between the world wars with the lack of a hegemon.[2]

This static perspective defines our starting point, on which we build to study countries' long-run dynamic incentives. This is critical, because the possibility of future changes in the structure of the world economy may induce a leading dominant country to keep an inherited rules-based system despite its short-run incentive to replace it with a power-based system. The reason is that reestablishing a rules-based system after it has been abandoned is costly, so keeping it today can help sustain it tomorrow. This may be helpful for the current leader if, in the future, it loses its dominance.

We analyze this problem in a many-country, infinite-horizon game where a stochastic process determines which country is the leader in each period. The current leader inherits the trade regime from the previous period and determines the current trade regime. Keeping the regime unchanged or moving from a rules-based to a power-based system is relatively

[1]This could occur, for example, when production requires irreversible investments, and the private sector allocates resources to export-oriented sectors in anticipation of a liberal trading environment (as in McLaren, 1997). This weakens the ex-post bargaining power of smaller countries to the point where, anticipating such outcomes, they are better-off opting out of power-based deals altogether.

[2]In Appendix A, we discuss the Hegemonic Stability Theory in more detail. We also highlight the common features of our model and of the HST, as well as their differences. A key distinction is that ours is a formal model of the long-run sustainability of trade regimes that explicitly considers dynamic trade-offs, whereas the HST is fundamentally a heuristic static theory. Our model extends the reach of the HST by formalizing some of its results, qualifying others, and significantly broadening its scope.

easy, but (re-)establishing a rules-based regime is costly for the leader, who must bear the coordination and implementation burden. The leader's choice will depend on whether it is hegemonic or simply dominant.

To assess the long-run viability of a rules-based trade regime, we use a theoretical framework of stochastic asynchronous games. Stochastic games model dynamic interactions in which the environment and underlying state evolve stochastically, possibly as a function of players' behavior (see Shapley, 1953). In asynchronous games, players act at different times rather than simultaneously. We focus on symmetric Markovian equilibria, which enable us to study how leading countries optimally choose their actions based on the current state of the economy and on expected future play, but not past play. We believe that this class of strategies captures well the strategic elements of such a long-horizon game.

We show that a hegemon that inherits a rules-based regime will always keep it. The reason is that, in addition to not having a short-run gain from replacing it (by assumption), there is no long-run gain either. But in any symmetric Markovian equilibrium in pure strategies, a non-hegemonic dominant country that inherits a power-based regime will always keep it as well. That is, a dominant country will not find it profitable to pay the cost of creating rules. This yields some bold lessons. In particular, it implies that once a power-based system is in place, the world can switch back to a rules-based regime only if it returns to a hegemonic state. Put differently, a rules-based trading system can emerge *only* when there is a hegemon to sponsor it. On the other hand, contrary to the claims of the HST, hegemons are not essential for *maintaining* a rules-based trading regime. With sufficient leadership turnover, dominant countries uphold such a regime to avoid entrenching a power-based system that would later disadvantage them.

We find that the general conditions that ensure a rules-based equilibrium are rather strict. In particular, we show that an equilibrium with a permanent rules-based regime can exist only if the cost to establish it is *not too small*. Intuitively, if the cost to establish a rules-based trading system is very small, then only the short run matters, since future leading countries can shift between regimes at virtually no cost. In that case, the trade regime loses its role as a state variable and the world will experience a cyclic equilibrium, transitioning between *rules* (under a hegemonic leader) and *power* (under a dominant leader) over time. Conversely, if the cost to establish a rules-based system is very large, then not even a hegemon would be willing to bear it, and the world will be permanently under a power-based regime. Thus, the existence and endurance of a rules-based system require the cost to reestablish it to be neither too large nor too small. Countries must also be sufficiently forward-looking, or else the equilibrium will be either cyclic or power-based. Furthermore, since the key reason dominant countries may uphold a rules-based regime is to avoid future losses once they become subordinate, a rules-based equilibrium also requires large enough turnover in world leadership.

We show as well that the decentralized equilibrium is often inefficient, as it fails to maximize the present value of the world's aggregate payoff. In particular, every cyclic equilibrium is wasteful, as the repeated transitions between rules-based and power-based regimes impose unnecessary costs. Additionally, many power-based equilibria are inefficient because dominant leaders do not fully internalize the losses they impose on other countries. In fact, an important class of inefficient power-based equilibria arises due to a collective action problem, in the sense that if a rules-based regime were implemented, it would persist. The problem is that no leading country may be willing to incur the initial cost of setting it up.

In the benchmark model, we assume that the states of the world evolve exogenously. We relax this assumption by allowing dominant leaders to extend their dominance when choosing a power-based regime. While this introduces an additional consideration, it does not alter the fundamental trade-offs of the benchmark model. Interestingly, this change may not make a power-based equilibrium more likely. Although the direct "endogenous dominance effect" surely makes a power-based regime more appealing for dominant countries, there are indirect effects that can offset it. By choosing *power*, a leader also raises the future cost of losing leadership: subsequent dominants become more likely to inherit and perpetuate *power*, thereby locking in a regime that will ultimately disadvantage today's leader once it becomes subordinate. This entrenchment effect can outweigh the temporary appeal of prolonging dominance, and thus may lead dominant countries to preserve a rules-based system despite their immediate incentives to abandon it.

In the discussion above, we consider that there is always a leading country that is decisive in determining the trade regime. Historically, this has often been the case, but there are moments when the world lacks a clear single leader. We consider two alternative scenarios.

First, we allow for a bipolar state in which the decision about the trade regime may depend on both (dominant) leaders. We consider that, in such a state, the two dominant countries play a simultaneous game to define the current trade regime, and the short-run payoffs have a prisoner's dilemma structure. This reflects the idea that a large country can benefit from exploiting its market power in international markets, and this benefit is greater when the other large country chooses trade barriers without attempting to exploit its own market power. In this context, we identify three new potential challenges to the sustainability of a rules-based trading system. One is the necessity to coordinate to sustain *rules*, if both leaders are needed to sustain it. We show that the scope for a rules-based equilibrium may decrease, but that, in general, coordination can still be achieved. Another issue is the desire to exploit market power. We show that this imposes greater difficulties for sustaining a rules-based equilibrium, relative to a world in which the likelihood of a bipolar state is minimal. A final issue is the possibility of free-riding. If one leader alone can sustain a rules-based regime, there will be an incentive to free-ride. We show that this tends to undermine the sustainability of a rules-based equilibrium. We conclude that sustaining a rules-based equilibrium in a bipolar state poses critical challenges, stemming from free-riding and market-power motives.[3]

Second, we introduce the possibility of a multipolar state, where leaders are "small." To capture the idea that, in the absence of a dominant leader, a cooperative regime requires coordination efforts from multiple players, we assume that establishing *rules* requires support from a group of countries, and each incurs a cost from doing so. Consistent with the literature, we assume that, net of the implementation cost, each country in the multipolar state prefers *rules* over *power*. We show that it is easier to sustain a rules-based equilibrium when the world has this multipolar state than when the world has a bipolar state. The main reason is that, in the former, market-power incentives are weaker.

WTO disciplines ("rules") are not immutable, however. Indeed, several scholars have proposed reforms (Grossman and Sykes, 2025; Mavroidis, 2025; Stiglitz and Rodrik, 2024). Our analysis provides the foundation for understanding the key trade-offs that any successful reform of the rules-based system would face. In particular, can the rules be adjusted to prevent a dominant leader from dismantling the entire system?

A first lesson is straightforward but important: if the system becomes more efficient and the resulting gains are shared between leading and non-leading countries, then sustaining it becomes easier. Indeed, studies such as Nicita et al. (2018) show that the scope for cooperation under the WTO is far from exhausted. If such an arrangement proves infeasible, another possibility is to exclude the disruptive dominant from the system. Replacing it with multipolar leadership could make the regime more resilient by reducing market power, but the lower gains from trade in such a "WTO-1" regime could undermine sustainability. A third possibility is a "rules-lite" system that offers additional gains to dominant countries, thereby increasing their short-run incentive to preserve it. We show that an "optimal appeasement" reform would provide the minimum necessary extra gains to leaders (while also minimizing the losses of non-leading countries) needed to induce them to keep the system in place.

Our study is motivated by Mattoo and Staiger's (2020) heuristic analysis of the possible motives behind the recent shift in U.S. trade policies, where they aim to understand why these actions are happening now and their consequences for the multilateral trading system.

[3]Although the settings are quite different – see footnote 7 – this finding resonates with the analysis of Becko et al. (2025) in their calibrated exercise, where they conclude that there is less global openness in a bipolar world than in a unipolar world.

Mattoo and Staiger argue that the U.S. has moved from rules-based to power-based tariff bargaining, and that this change is related to the U.S. loss of hegemonic power.[4] They warn that this shift poses a threat to the sustainability of the rules-based trading system, and risk pushing it to its breaking point. In doing so, the authors bring up a discussion about the long-run implications for the sustainability of the rules-based system associated with changing hegemonic power. However, they do not develop a formal theoretical analysis to address these scenarios. This is precisely what we do in this paper. Specifically, we rely on their insights to motivate some key building blocks of our model. We then formally investigate the long-run implications of changing leadership and agenda-setting power in the international trading system, providing a framework for understanding its evolving landscape. To our knowledge, our dynamic analysis is entirely novel in both the trade policy and international relations literatures.[5] Given the fizzling multilateral cooperation observed in recent years and the current uncertainties about the sustainability of a rules-based trade regime, it is also very topical.

Most of our results are valid regardless of the ranking of aggregate welfare under rules- and power-based regimes, but for some results we consider the world to be better-off under a rules-based regime, an assumption that is grounded on established research. Bagwell and Staiger (1999) analyze tariff negotiations under the institutional setting of the General Agreement on Tariffs and Trade (GATT) and show that the principles of reciprocity and nondiscrimination can guide governments in implementing trade agreements that are immune from the terms-of-trade inefficiency that arises with unilateral trade policies, while mitigating the effects of bargaining power on negotiations.[6] Quantitative models have provided a better understanding of the magnitudes involved. Ossa (2014) shows that cooperative tariffs, obtained through multilateral trade negotiations, yield large welfare gains. Naturally, a rules-based system has value only if there is a mechanism to enforce the rules. The

[4]Zissimos (2022) links this change in the U.S. support for a rules-based system to the rise of populism.

[5]In a recent, complementary study, Broner et al. (2025) provide a novel model relating the presence of a hegemon and the degree of globalization, relying on a trade-off between gains from trade stemming from similar policies and country heterogeneity over those policies. Related to our paper, they study how hegemons affect economic integration. However, our model differs from theirs in the questions addressed, the modeling approach, and, consequently, in the insights obtained.

[6]Several studies provide empirical support for the claim that the GATT/WTO key principles encourage multilateral trade liberalization. For instance, Ludema and Mayda (2013) develop and estimate a model showing that trade agreements based on GATT/WTO rules can neutralize the terms-of-trade externality and thus induce lower average applied tariffs. Bagwell et al. (2021) develop and estimate a model to evaluate the impacts of the principle of nondiscrimination. Through counterfactual analysis, they find that tariff negotiations during GATT's Uruguay Round resulted in increased global welfare compared to counterfactual discriminatory negotiations, and that nondiscrimination was particularly beneficial for countries in weak bargaining positions. Broner et al. (2025) show, using a sample spanning 180 years, that hegemons are associated with more – and more effective – trade agreements.

GATT/WTO system has an explicit mechanism for dispute resolution. Maggi (1999) develops a model that emphasizes the roles played by this mechanism in promoting multilateral trade negotiations and in informing third countries about violations of the agreement.

At a broader level, our paper also relates to the burgeoning research in geoeconomics (Clayton et al., forthcoming). A key contribution of our analysis is the incorporation of an explicit dynamic component, which is essential for understanding the long-term prospects of the rules-based world trading system. At the same time, we abstract from other important issues that could be related to the trade regime, such as geopolitical rivalry due to countries' relative economic positions (Mattoo et al., 2025); concerns over economic security (Clayton et al.; 2024); economic dependence shaping political alignment (Kleinman et al., 2024); and countries' actions to shape policy beyond their borders (Broner et al., 2025; Camboni and Porcellacchia, 2025).

We also abstract from the possibility that a subset of countries could form exclusive rules-based agreements – like the free trade agreements that have sprung out in the last 30 years (for surveys of that literature, see Freund and Ornelas, 2010; Limao, 2016; Maggi, 2014). Historically, these agreements have operated within the WTO framework, which allows their formation under specific conditions. By contrast, a world where multilateral trade negotiations are governed by power while bilateral trade agreements are based on rules would be significantly different – it may resemble, in particular, the trading environment of the 1930s.[7]

The remainder of the paper is organized as follows. In section 2, we describe the origins and workings of the GATT/WTO rules-based system, its current challenges, and some worldwide changes in economic power since 1900. In section 3, we develop the benchmark model. In section 4, we endogenize the transition probabilities. In section 5, we extend the model to allow for a state where there are two leading countries, whereas in section 6, we allow for a state where there is no leading country. We discuss the trade-offs behind a possible reform of the rules-based system in section 7. We conclude in section 8.

2 Institutional Background

2.1 The GATT/WTO Rules-Based Multilateral Trading System

The years of instability after World War I, followed by the Great Depression, explain the main motivation for establishing a formal system of international trade rules. The worldwide

[7] Becko et al. (2025) study the related situation in which a hegemon (or two competing hegemons) chooses optimal tariffs based on terms-of-trade considerations but is also motivated by the desire to align smaller countries geopolitically, which they do by offering participation in FTAs.

economy during that time was characterized by a drop in the volume of trade, a decline in world industrial production, and widespread unemployment. As economic conditions worsened, countries started to carry out protectionist policies in the form of tariffs and non-tariff barriers. The international trading system was fragmented, with trade policy decisions taken unilaterally in a noncooperative environment, governed mainly by bilateral agreements. As a result, trade barriers became increasingly restrictive and followed "beggar-my-neighbor" motivations (Irwin, 2020; Staiger, 2022). According to Staiger (2022), the main inefficiency of this setup was the disregard of the negative international externalities of trade policies on trade partners and third parties.

Against this backdrop, the establishment of the GATT in 1948 represented a significant shift in the governance of international trade. In 1995, the GATT was subsumed by the newly created WTO, which built upon the economic and legal principles of its predecessor. Arguably, through rounds of trade negotiations under a set of agreed-upon rules, the GATT/WTO system has laid the ground for the progressive reduction of trade barriers among its members. As Staiger (2022) points out, the rules "simplify the tariff bargaining problem and make it manageable, and this can help countries negotiate more efficient policies" (p. 233). Indeed, since 1948, industrialized countries have reduced average tariffs on manufactured goods from about 40 percent to less than 5 percent in 2024. Furthermore, numerous developing economies have joined the system over the years.

Several authors have argued that these achievements were made possible because of the rules of the GATT/WTO, which define the current multilateral trading system. The system is built on the principles of reciprocity and nondiscrimination, together with a dispute settlement procedure (DSP) to discipline and resolve trade disputes between member countries. These correspond to the legal foundations of global trade that guarantee trade rights to all WTO members while restricting the exercise of power by individual countries.

As Bagwell and Staiger (1999) put it, the principle of reciprocity is a "GATT norm under which one country agrees to reduce its level of protection in return for a reciprocal 'concession' from its trading partner" (p. 217). In other words, the principle specifies that bargains and concessions should be balanced. For this reason, it serves to mitigate the exercise of power in tariff bargains. Thus, if a country takes the unilateral decision to increase its tariffs, its partners could follow the same strategy, undermining any advantage gained by the country that had moved first.

In turn, nondiscrimination specifies that, every time a country changes a trade barrier on a given product or service applied to the imports of one trading partner, it must extend the policy change to all other WTO members. This ensures that any export gains derived from a lower foreign tariff will be shared with third countries. It therefore mitigates a powerful country's ability to capture the gains from bargaining down the tariffs of a weaker partner, since any tariff changes will be extended to all partners, whether rich or poor, weak or strong.[8]

Finally, the WTO's DSP is a mechanism for resolving trade conflicts that governments access when they consider that their rights under the GATT/WTO agreements are being violated. According to Maggi (1999), the DSP also has an informational role as it identifies agreements violations and informs third countries, exposing the offending country to a loss of reputation in the trading community. Thus, even though the GATT/WTO system lacks the power to directly enforce trade agreements, the DSP, together with the principles of reciprocity and nondiscrimination, are mechanisms for facilitating and enforcing the rules of the multilateral trading system.[9]

2.2 The Erosion of the Rules-Based System

Despite its historical success, the GATT/WTO trading system has faced a series of challenges in recent decades that may threaten its sustainability. Irwin (2020) argues that the main concern is the impact of the recent trade wars, primarily waged between the United States and China, on the rules-based trade regime. These tensions seem unlikely to recede in the near future. In this section, we provide evidence of the erosion of the rules-based system in the last 15 years, but without including U.S.-led developments of 2025, which follow from, but drastically escalate, earlier developments.

The following two figures illustrate the increasing disregard for the GATT/WTO system since 2008, using information from the Global Trade Alert (GTA) database. The GTA documents unilateral government statements concerning foreign commercial interests, which encompass actions related to trade in goods and services, investment, and labor force migration. The database evaluates the types of interventions resulting from each government act. These interventions can be either beneficial or harmful depending on the relative treatment of foreign versus domestic commercial interests.[10]

[8]Bagwell and Staiger (2004, 2010, 2018) provide further theoretical treatments for the benefits of the reciprocity and nondiscrimination principles. Bagwell et al. (2020) provide empirical evidence based on detailed tariff bargaining data from an early GATT round of negotiations.

[9]Because it provides policy stability, a rules-based system also has the benefit of increasing supply chain resilience. As Ossa (2023) argues, "the rules-based nature of the multilateral trading system is particularly important for supply chain security [but] these advantages would be lost in a power-based trading system" (p. 30).

[10]The GTA uses three different classifications: (i) the intervention almost certainly discriminates against foreign commercial interests; (ii) the intervention likely involves discrimination against foreign commercial interests; and (iii) the intervention liberalizes on a non-discriminatory basis or improves the transparency of a relevant policy. In our data, we merge the first and the second classifications into the category of "protecting," while denoting the third as "liberalizing." In the database, a single state act can result in one

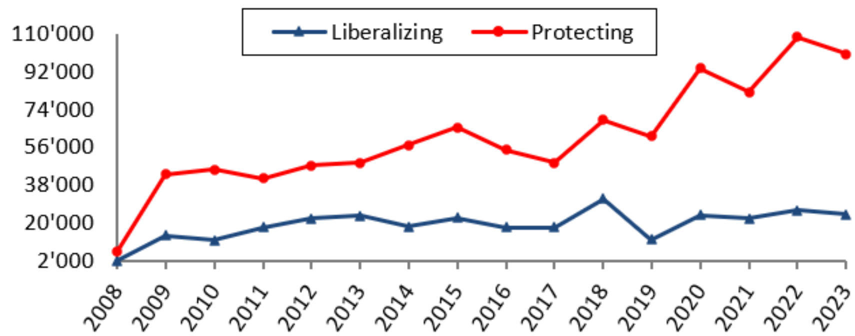

Figure 1: World trade policy interventions (2008 - 2023)
Source: Global Trade Alert database, 2024

Figure 1 shows the number of liberalizing and protecting measures since 2008. The latter has been consistently higher than the former, by a large margin. Moreover, the difference has widened since 2017, as the number of protecting measures kept increasing, while the number of liberalizing interventions has been rather stable.

Many of the interventions in Figure 1 reflect the ongoing trade war between the U.S. and China. Their policies also draw more interest by themselves, since they are the two largest trading nations. Figure 2 then considers the same data as in Figure 1, but only for those two countries. The difference between protecting and liberalizing is even starker in that case. The number of liberalizing measures in the U.S. and China since 2008 has been not only flat, but also tiny. In contrast, the number of protecting measures has been significantly higher, and shows a sharp increase in recent years.

It is important to note that, while the trend toward protectionism is more visible for the two main economies, it is more general. In particular, if we subtract American and Chinese measures, the curves in Figure 1 would shift down but still display the same qualitative pattern. In other words, the trend toward more protectionist measures in the last 15 years is accentuated by the behavior of the two main economic powers, but is not limited to them. We replicate this figure, excluding data from the U.S. and China, in Appendix B.

Indeed, using data from 23 developed economies, Colantone et al. (2022) document a persistent shift of voters toward more isolationist views since the early 1990s, and especially since the financial crisis. They show that this change in voters' preferences has been accompanied by analogous changes in government positioning. Goldberg and Reed (2023) make a similar point.

or more interventions, and may be beneficial or harmful to multiple countries simultaneously. In our data analysis, we consider the number of interventions resulting from an act. For instance, if an act impacts twenty countries, it is counted as twenty interventions. This has the advantage of assigning greater importance to measures that affect more countries. Alternatively, we could consider only the number of acts. The number of measures would be lower, but the resulting pattern would be very similar.

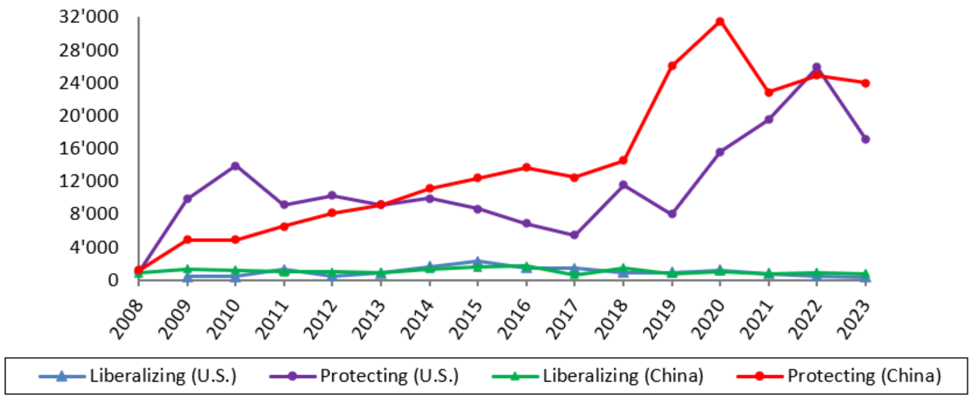

Figure 2: U.S. and China trade policy interventions (2008 - 2023)
Source: Global Trade Alert database, 2024

We also observe increasing hostility in trade disputes. While it is not easy to formally document such hostility, the data from WTO disputes provide an angle to this issue. When WTO members believe that their rights are being infringed with respect to any agreement under the GATT/WTO rules, they can request consultations to initiate a dispute under the Dispute Settlement Understanding. This marks the first stage in the dispute settlement process. Sometimes, the parties involved in a dispute resolve their disagreements through consultations. But if they are unable to reach an agreement through consultations, the complainant can request the establishment of a panel, which represents the second stage of the process. Figure 3 shows the share of disputes transitioning from consultations to panels in the quadrenniums since the creation of the WTO in 1995. We observe a clear increase in this share since 2007-2010. This trend may signal a growing difficulty for WTO members in resolving disputes solely through the consultation mechanism.[11]

Now, after a panel issues a verdict on a case, any of the parties can take the matter to the WTO's Appellate Body. In the Appellate Body there are nine judges with fixed mandates, and it requires a minimum of three judges to deliberate over each case. Since the Obama Administration, the U.S. has blocked all new appointees. As a result, the number of members of the Appellate Body has been decreasing as their mandates expire. Indeed, since 2019, the Appellate Body has had fewer than three members (it currently has none), thereby becoming unable to consider any cases. This, of course, drastically reduces the value of bringing a case to the WTO in the first place, as the formal enforcement mechanism has become literally inoperative.

[11] This is also the conclusion of Mavroidis (2022), who offers a detailed analysis of the WTO's disputes since its inception. A related sign of hostility and disregard for the WTO system is the recent, unprecedented tendency of members (initiated by the U.S.) to invoke national security concerns when imposing measures that violate their existing commitments. See Mavroidis (2025) for an in-depth discussion of the complexities and implications of this behavior.

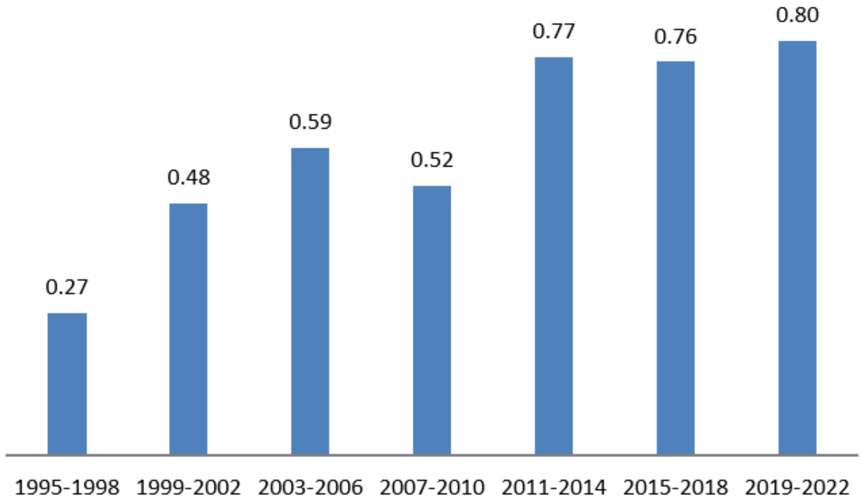

Figure 3: Share of DSP consultations moved to panels (1995 - 2022)

Source: WTO Annual Report, 2023

Staiger (2022) highlights other challenges for the rules-based multilateral trading system (e.g., the rise of large emerging economies, digital trade, and the proliferation of global value chains). Yet the primary challenge he focuses on, which provides the background for our theoretical model, is the decline of U.S. hegemonic power and its potential implications for the sustainability of the world trading system.

2.3 A Fading Hegemon

By the end of World War II, the U.S. had become an international hegemonic power, and its leaders have strongly supported and committed to the rules-based regime that has guided the international trading system since then. This hegemony lasted uncontested for several decades, but the country faced a slow but steady decline in its economic power over that period. In parallel, in the last decades, the world has witnessed the rise of China, which became the second-largest economy in the world in 2010. In that sense, China appears to pose a challenge to the U.S. for global dominance.

Figure 4 shows the evolution of the U.S. and China's share of world GDP at current prices since 1900. The U.S. emerged as the dominant global economic power in the early 20th century, a position it consolidated after World War II, when it reached a world share of GDP above 45%. This post-war peak coincides with the creation of the GATT in 1947 (represented by the vertical green line), which institutionalized American-led trade liberalization. However, from the late 1960s onward, the U.S. share of global GDP shows a steady decline.

Meanwhile, China's trajectory follows a different path. After maintaining a relatively low and decreasing share of global GDP until the early 1980s, it began to show a rapid rise in the late 20th century. This growth accelerated significantly after China joined the WTO in 2001, marking its deeper integration into global markets. The gap between the two economies has narrowed significantly since then, with the U.S. share of world GDP at 26% and China's at 17% in 2023. Although the U.S. remains the leading economic force, the rise of China implies that it no longer holds the unchallenged hegemony it once did.

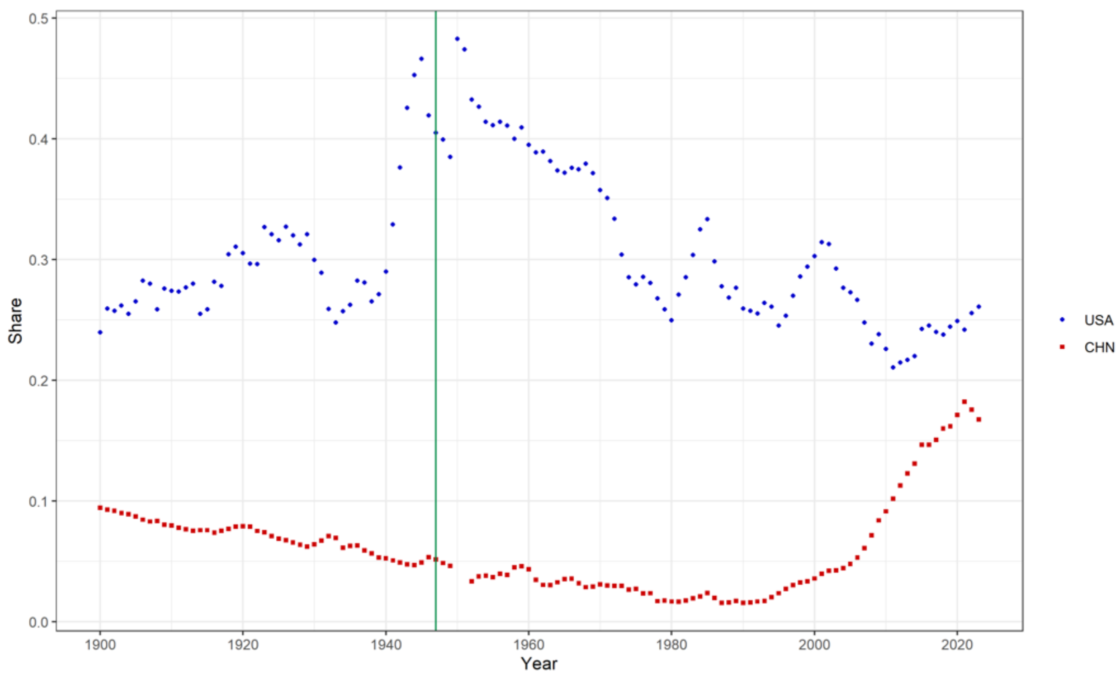

Figure 4:
Share of world GDP (1900 - 2023, Current Prices)

Source: 1900-1949 data from Klasing and Milionis, 2014; 1950-1959 data from Penn World Tables 10.01 (Feenstra et al., 2015); 1960-2023 data from the World Bank's World Development Indicators, 2025 (https://databank.worldbank.org/)

Mattoo and Staiger (2020) argue that, with the decline of its relative economic importance, the U.S. no longer perceives the benefits of committing to a system of rules that restricts its power. As a result, it has been moving away from "rules-based" tariff negotiations to engage in "power-based" negotiations. The purpose of this practice is to make its partners reduce their bilateral tariffs vis-à-vis the U.S., in a way that departs from the principles of reciprocity and nondiscrimination as described in the GATT/WTO, and instead rests on countries' bargaining power.[12, 13]

Will those changes undermine the long-run sustainability of the rules-based multilateral trading system, or could they be regarded as transitory disturbances? Our model answers this and other questions.

3 Benchmark Model

3.1 Preliminaries

We consider an environment where there are N countries and $2N$ states of the world. Denote the set of countries by $\mathcal{I}= \{1, 2, ..., N\}$, and the set of states of the world by $\Omega = \{h_1, d_1, h_2, d_2, ..., h_N, d_N\}$. State h_i represents the state where country i is a *hegemon*, and d_i represents the state where country i is *dominant*. We denote by ω a typical element of Ω.

When a country is either a hegemon or a dominant, we call it the *leader*. As such, it chooses the trade regime for the current period.[14] Specifically, in states h_i and d_i country i chooses a trade regime ρ from the set $\{P, R\}$, where P stands for "power-based" and R stands for "rules-based." When a country is neither a hegemon nor a dominant, we say it is *subordinate*. As such, it does not take any action concerning the trade regime in those states. At any state, one country is the leader (either hegemon or dominant) and $N - 1$ countries are subordinate.

Let $u_i(\omega, \rho)$ denote the stage-game payoff of country i when the state of the world is $\omega \in \Omega$ and the leader chooses regime $\rho \in \{P, R\}$. Our parsimonious framework relies solely

[12] As Horn and Mavroidis (2025) observe, the U.S. has challenged the multilateral rules before. The first systematic violation occurred under the Reagan Administration – when the U.S. share of world GDP reached a local trough and Japan was perceived as a challenger to America's dominance. Subsequent violations have taken place since the George W. Bush Administration; however, "the current scale of deviation from multilateral rules is simply unprecedented" (Horn and Mavroidis, 2025).

[13] As we briefly discuss in Appendix C, hegemonic Britain in the 19th century also promoted freer trade following rules-based negotiations, although the rules were surely less sophisticated than the GATT. And yet, as Britain started to lose its hegemony late in that century, it shifted to power-based, rules-less practices.

[14] As Norrlof (2015) defines, "in international relations, hegemony refers to the ability of an actor with overwhelming capability to shape the international system." Here, 'to shape the international system' means choosing the trading regime for the current period. We give this prerogative to the leader, be it hegemon or 'simply' dominant. This is similar to the assumption used by Aghion et al. (2007), where the leading country decides between multilateral and sequential bargaining when negotiating preferential trade agreements. It is also in the same spirit as the approach of Farhi and Maggiori (2018), who study the properties of the international monetary system with and without a hegemon that is the sole issuer of international reserve assets.

on the following assumptions on stage-game payoffs:

$$\begin{aligned} &u_i(\omega, R) > u_i(\omega, P),\ \forall \omega \in \Omega \setminus \{d_i\}, \forall i; \\ &u_i(d_i, P) > u_i(d_i, R),\ \forall i. \end{aligned} \tag{1}$$

The first set of inequalities indicates that when i is hegemonic or subordinate, its static gains are higher under regime R than under regime P. In turn, the second inequality says that when i is dominant, its static gains are higher under regime P.

The rationalization of this ordering follows Mattoo and Staiger (2020), who provide a detailed micro-founded discussion, based on an extensive literature, of how countries would fare under different trade regimes depending on their economic sizes. The mitigation of the power of larger countries in tariff bargaining due to the reciprocity and nondiscrimination rules encourages the participation of weaker countries in the system. Without those rules, weaker countries may fear excessive exploitation and choose to avoid any type of bargaining. This scenario could arise in the presence of sunk investments, as shown by McLaren (1997) and discussed by Bagwell and Staiger (1999). Private agents invest in the exporting sector and divest in the import-competing sector in the expectation of bilateral tariff reductions vis-à-vis a large trading partner. After investments are sunk, the bargaining position of the large country is bolstered by the economic specialization of the smaller countries. The large country would therefore have an incentive to "hold up" the smaller countries and renegotiate the original agreement in a way that benefits itself at their expense. If the large country is too powerful, this would push the smaller countries to a situation that leaves them worse off than they would have been in the absence of any agreement. Anticipating this scenario, negotiations may never get off the ground. In a different context, Clayton et al. (2024) also build on the idea that a hegemon can improve its own welfare by tying its own hands through a commitment device, like the creation of an international organization such as the WTO.[15]

[15] In a similar vein, Goldstein and Gowa (2002) argue that one of the reasons why the U.S. encouraged the formation of the current multilateral trading system was the existence of large power asymmetries in the postwar period. They argue that the rules-based system allowed weaker countries to effectively discipline the behavior of the U.S. in trading relations, and that empowerment encouraged their participation in the system. In that sense, the establishment of rules in the world trading system can be interpreted as a commitment technology against future expropriation of small countries. The assumption that rules help weak countries relative to strong countries in tariff negotiations is supported by the empirical analysis of Bagwell et al. (2021). This rationale parallels the reasoning developed by Maggi and Rodriguez-Clare (1998, 2007) – whereby trade agreements are used as a *domestic* commitment device against lobbying by firms – to an international environment. In a broader sense, it also relates to the logic developed by Acemoglu and Robinson (2000) to explain why elites in Western countries voluntarily extended voting rights as a commitment against future expropriations. Unlike those papers, here the commitment (offered by a rules-based regime) is useful in a static sense, but not in the long run.

We label a country *hegemon* when it is large enough to generate this reaction by small countries. In this case, the large country would be better off under a rules-based regime, because it provides insurance to weaker countries that such exploitative renegotiation of previous agreements will not occur. Naturally, this tends to boost participation – much like the GATT/WTO system has experienced over more than seven decades. Accordingly, both small countries and the hegemon tend to benefit from a rules-based system that mitigates the role of power in negotiations. Arguably, this is what compelled the U.S. to sponsor the post-WWII international order, when its relative economic power was at its historical peak.

The situation is different when power asymmetry is not so significant, so that weaker countries would engage in tariff negotiations even in the absence of rules. We label a country *dominant* when it generates this reaction by small countries. The dominant country then benefits from exploiting its stronger bargaining power because that would not preclude participation of subordinate countries. In Appendix D, we elaborate further on the ordering of payoffs under each state.

Hence, *from a static point of view*, we have that:

- *Subordinate* countries prefer R to P, because R protects them from exploitation in tariff negotiations;

- *Dominant* countries prefer P to R, because they can benefit from their privileged bargaining power in tariff negotiations;

- *Hegemon* countries prefer R to P, because R serves as a commitment device to induce participation of subordinate countries in tariff negotiations.

It is worth pausing briefly to discuss the assumption that hegemons prefer *rules* over *power*, net of the cost c, in a static sense. The assumption will be important to justify the formation of a rules-based regime, consistent with the actions of the United States after World War II and the United Kingdom in the mid-19th century. However, as will become clear later, this assumption is not required to explain the persistence of such a regime. We will return to this point in section 3.5, when we discuss equilibria when hegemons prefer P to R in a static sense.

3.2 Dynamic Model

We now construct a dynamic model of trade regimes to study the conditions for the sustainability of a rules-based system when dominant countries have an intrinsic incentive to favor a power-based system.

The initial state and the initial trade regime, at period $t = 0$, are exogenously set to (ω_0, ρ_0). Since we are concerned with a model that can account for the creation of the rules-based regime, we consider the case in which $\rho_0 = P$.[16] For concreteness, we set $\omega_0 = h_1$.

From period $t = 1$ onward, the transition between states is governed by a Markov matrix Q, in which element $q_{\omega,\omega'}$ is the probability of moving from state ω to state ω'. We impose only two restrictions on these probabilities: (i) there is no absorbing state – i.e., $q_{\omega,\omega} < 1$ for all $\omega \in \Omega$; and (ii) the transition probabilities from a hegemonic state to another where a different country becomes the leader, and from a dominant state to another where a different country becomes the hegemon, are zero: $q_{d_i,h_j} = q_{h_i,d_j} = q_{h_i,h_j} = 0$ for all $i, j \in \{1, 2, ..., N\}$ with $i \neq j$. The first assumption implies simply that there is no "final state" of the world. The second assumption captures the idea that the transition to and from a hegemonic state is "smooth," in the sense that a country becomes dominant before (possibly) turning hegemonic, and a hegemon first becomes dominant before (possibly) losing its leadership. More precisely, if a country is a hegemon in period t, in period $t + 1$ either it remains a hegemon or becomes dominant; if a country is dominant in t, in $t + 1$ it remains dominant, becomes hegemonic, or another country becomes dominant. Figure 5 presents a visualization of this transition rule.

The matrix Q induces a unique stationary probability distribution over the states in Ω. Let $\mu_t \in \Delta(\Omega)$ denote the probability distribution over states at time t, with an initial distribution $\mu_0 = (1, 0, 0, \ldots, 0)$. The distribution evolves according to $\mu_{t+1} = \mu_t Q$, so that in general $\mu_t = \mu_0 Q^t$. The stationary distribution, denoted by μ, is given by $\mu \equiv \lim_{t\to\infty} \mu_1 Q^t$, and can equivalently be found by solving $\mu = \mu Q$. In the benchmark model, we take matrix Q as exogenous. In section 4, we allow the choice of regime to affect the transition probabilities.

Once the leading country chooses a trade regime, the subsequent period starts with that regime as the status quo. If the status quo regime is P, moving to R implies a cost $c > 0$ to the country making the change. On the other hand, moving from R to P is costless. Keeping the regime unchanged is also costless. The rationale is that putting in place a rules-based system is more costly because it requires establishing a series of institutional rules and procedures and putting together a bureaucracy to set and enforce them, as in the WTO, in addition to coordinating with the other countries to bring them to participate in the system. Conversely, as argued by Keohane (1984), keeping it in place is significantly less costly once it is already established. Similarly, reversing to/maintaining a power-based system is relatively less costly, as it does not require an institutional structure to guide and

[16] The case in which $\rho_0 = R$ is qualitatively similar, the main difference being that this case allows for the possibility of an equilibrium in which rules are always chosen on the equilibrium path even if no country is willing to incur the cost to introduce it.

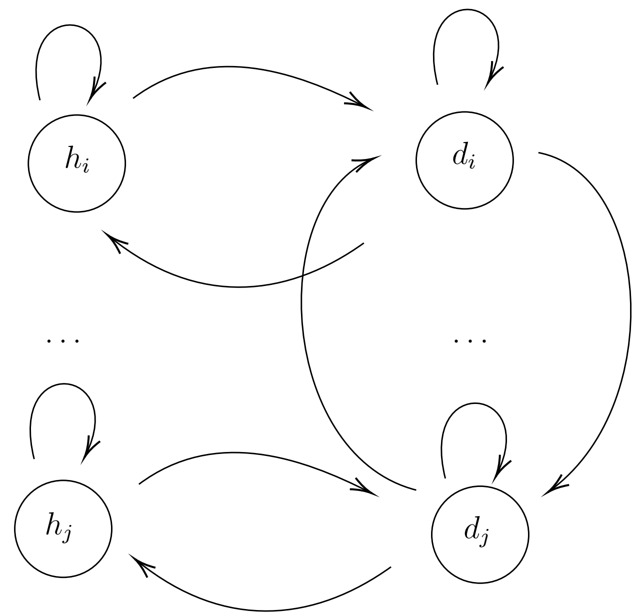

Figure 5: Transition between states

enforce it, or any multilateral coordination.[17]

We focus on symmetric Markovian equilibria, i.e., strategy profiles that induce subgame perfect Nash equilibria and that are Markovian and symmetric. This class of equilibria is intended to capture equilibria in which countries that find themselves in a leadership position under a particular trade regime will behave similarly. Formally, we define:

Definition 1 (Symmetric Markovian Strategies) *A Markovian strategy profile is described by $\sigma : \Omega \times \{P, R\} \mapsto \{P, R\}$, where $\sigma(\omega, \rho)$ is an element of that strategy that specifies a regime choice in state ω when the status quo regime is ρ. A strategy profile is symmetric if $\sigma(h_i, \rho) = \sigma(h_j, \rho)$ and $\sigma(d_i, \rho) = \sigma(d_j, \rho)$, $\forall\ i \neq j$ and $\forall\ \rho \in \{P, R\}$.*

We define symmetric Markovian equilibria as:

Definition 2 (Symmetric Markovian Equilibria) *A strategy profile is a symmetric Markovian equilibrium if it induces a subgame perfect Nash equilibrium and strategies are symmetric and Markovian.*

[17] Assuming that the regime transitions $\{R \to P\}$, $\{R \to R\}$ and $\{P \to P\}$ imply the same cost, as we do, is with loss of generality. However, further assuming that the common cost is zero, as we also do, is without additional loss of generality. That is, if we assumed instead that the regime transitions $\{R \to P\}$, $\{R \to R\}$ and $\{P \to P\}$ implied cost $c' \in (0, c)$, our analysis would remain unchanged. The reason is that all that matters for the regime choice is the difference in the cost of adopting one regime relative to the other, not their absolute cost.

Since we focus on symmetric Markovian equilibria, which are stationary by definition, we omit time subscripts in our exposition.

We now describe the recursive equations representing country i's payoff functions. We denote by $V_i(\omega, \rho)$ country i's expected discounted continuation payoff at state $\omega \in \Omega$ and received status quo trade regime $\rho \in \{P, R\}$. For $\omega \in \{h_i, d_i\}$, country i's payoff is:

$$V_i(\omega,\rho) = \max_{\{P,R\}}\{u_i(\omega,P) + \delta \sum_{\omega' \in \Omega} q_{\omega,\omega'} V_i(\omega',P), u_i(\omega,R) - \mathbb{1}_{\{\rho=P\}} c + \delta \sum_{\omega' \in \Omega} q_{\omega,\omega'} V_i(\omega',R)\}. \tag{2}$$

Essentially, the leader compares the discounted payoffs of choosing P or R today, taking into account that switching from P to R has cost c, and that today's regime choice affects the regime choice in the next period.

In contrast, at any state h_j or d_j, $\forall j \neq i$, country i is subordinate and does not choose the regime, so its value functions at those states depend on the continuation strategies and the regime choice of their opponents at that period. For $\omega \in \Omega \setminus \{h_i, d_i\}$, country i's payoff function is:

$$V_i(\omega,\rho) = \begin{cases} u_i(\omega,P) + \delta \sum_{\omega' \in \Omega} q_{\omega,\omega'} V_i(\omega',P) & \text{if the leader chooses } P \text{ at } (\omega,\rho), \\ u_i(\omega,R) + \delta \sum_{\omega' \in \Omega} q_{\omega,\omega'} V_i(\omega',R) & \text{if the leader chooses } R \text{ at } (\omega,\rho). \end{cases} \tag{3}$$

It will be convenient to define three important classes of symmetric Markovian equilibria in pure strategies:

Definition 3 (Classes of Equilibria) *A symmetric Markovian equilibrium in pure strategies is denoted by:*

1. *Rules-based equilibria: hegemon and dominant countries choose rules on the equilibrium path* $[\sigma(d_i, R) = \sigma(h_i, R) = \sigma(h_i, P) = R,$ *and* $\sigma(d_i, P) = P, \ \forall i \in \mathcal{I}]$.

2. *Power-based equilibria: hegemon and dominant countries choose power on the equilibrium path* $[\sigma(d_i, P) = \sigma(h_i, P) = P,$ *and* $\sigma(h_i, R) = R, \ \sigma(d_i, R) = \rho,$ *with* $\rho \in \{P, R\}, \ \forall i \in \mathcal{I}]$.

3. *Cyclic equilibria: hegemon countries choose rules and dominant countries choose power on the equilibrium path* $[\sigma(d_i, \rho) = P, \ \sigma(h_i, \rho) = R, \forall i \in \mathcal{I}, \ \forall \rho \in \{P, R\}]$.

Note that, in the strategies above, we also specify the off-equilibrium choices. The power-based equilibria specify the play of P on the equilibrium path, but we allow for two distinct groups of power-based equilibria, depending on the play at state (d_i, R), $\forall i \in \mathcal{I}$. We discuss these two subclasses of power-based strategies in section 3.5. Later in this section, it will become clear why the off-equilibrium choices must be the ones specified above.

Because we impose very few restrictions on the transition matrix Q and on the payoffs (equation (1)), establishing existence of pure-strategy symmetric Markovian equilibrium generally is difficult. However, we can show that, for small and large values of c, there is always an equilibrium. All proofs are in Appendix E.

Proposition 1 (Existence of Equilibrium for Small and Large Values of c) *Fix any given discount factor $\tilde{\delta} < 1$. Then, we have that:*

A. For c sufficiently small, a cyclic equilibrium always exists (and it is the unique pure strategy symmetric Markovian equilibrium). Moreover, for such values of c, a cyclic equilibrium will also exist for $\forall \delta < \tilde{\delta}$.

B. For c sufficiently large, there is always a power-based equilibrium (and it is the unique pure strategy symmetric Markovian equilibrium). Moreover, for such values of c, a power-based equilibrium will also exist for $\forall \delta < \tilde{\delta}$.

Intuitively, if the cost to establish a rules-based trading regime is very small, then only the short run matters, precisely because switching across regimes requires little cost. Therefore, with c small, when the world has a dominant leader (which prefers P in the short run), the regime is P; when the world has a hegemon leader (which prefers R in the short run), the regime is R. If instead c is very large, since the initial regime is P, the world will never get away from P, because not even a hegemon will be willing to incur the cost to change the regime. Furthermore, a rules-based equilibrium can only exist insofar agents are not too short-sighted. The upshot is that there are strict limits on the conditions under which rules-based equilibria can exist, as the following proposition shows.

Proposition 2 (Necessary Conditions for Rules-Based Equilibria) *The following conditions are necessary for the existence of rules-based equilibria:*

A. There exists $\underline{c} > 0$ such that a rules-based equilibrium can exist only if $c \geq \underline{c}$.

B. There exists $\bar{c} > \underline{c}$ such that a rules-based equilibrium can exist only if $c \leq \bar{c}$.

C. There exists $\underline{\delta} > 0$ such that a rules-based equilibrium can exist only if $\delta \geq \underline{\delta}$.

Therefore, a sufficiently high cost to (re-)establish R is needed to make the trade regime a relevant state variable, where a dominant country chooses to keep R in the expectation that R will be maintained when it is no longer dominant. Thus, an equilibrium with a permanent rules-based regime can exist only if the cost to establish it is *not too small*. However, the cost of establishing a rules-based regime cannot be too large either, as no country would ever pay for it, and the system would remain indefinitely at the initial status quo, P. Finally, countries must be sufficiently forward-looking so that dominant leaders perceive a long-term benefit in maintaining regime R despite it not being their most advantageous static choice.[18]

Now, to further characterize equilibrium trade regimes, we need to understand how dynamic considerations shape players' incentives. We first show that it is always weakly better to inherit status quo regime R rather than P.

Lemma 1 (Rules-Based Regimes are Preferred as Status Quo) *In any symmetric Markovian equilibrium, every country weakly prefers to enter every state under regime R:*

$$V_i(\omega, R) \geq V_i(\omega, P), \ \forall \omega \in \Omega, \ \forall i \in \{1, 2, ..., N\}.$$

This result follows from two observations. First, whenever country i is the leader, switching to P is costless, unlike switching from P to R. Second, whenever country i is subordinate, it weakly prefers that the leader starts at regime R, because the cost c may discourage a leader from switching from regime P to R (recall that R would be beneficial to all subordinate countries), but the opposite does not happen.

Lemma 1 may be interpreted as suggesting that it may be "easy" to sustain regime R. As we will see, that interpretation is not correct. A key result is that, in any symmetric Markovian equilibrium, when inheriting their myopically optimal status quo, leading countries always choose to keep it.

Proposition 3 ***(No Hard Choices: Countries Keep Myopically Optimal Status Quo)*** *In any pure strategy symmetric Markovian equilibrium,* $\sigma(h_i, R) = R$ *and* $\sigma(d_i, P) = P$*,* $\forall i \in \{1, 2, ..., N\}$.

The first part of Proposition 3 holds because, when there is a hegemon, if it inherits a rules-based regime, there is no short-run gain from replacing it, and there is no long-run gain either. The latter follows because it is always weakly better to enter a state under regime R than under regime P (Lemma 1), so leaving R for the next period is preferred to leaving P.

To see why the second part of Proposition 3 holds, suppose, by contradiction, that $\sigma(d_j, P) = R$ instead of $\sigma(d_j, P) = P$. First, note that when a dominant country inherits a power-based regime, in the short run it is better off maintaining P, by assumption.

[18] For clarity, we assume that the discount factor δ is common across countries, but it would be straightforward to consider the situation where countries are heterogeneous in that dimension. In this case, the condition in Proposition 2(c) would need to be satisfied for all countries that could at some point become leaders. Therefore, the last inequality would become $min_i\{\delta_i\} \geq \underline{\delta}$.

Second, the dominant leader has no long-run gains from switching to R either. To see that, consider all transition possibilities in the subsequent period. The possibilities are the following: the dominant country j may (i) remain dominant, (ii) become hegemon, or (iii) another country may become dominant. If j remains dominant in the subsequent period, it would have benefited from keeping P. If j becomes hegemon, it would again have benefited from keeping P (and potentially choosing R only after becoming a hegemon). Finally, if country j becomes subordinate, then it would have been better off maintaining P, since the new dominant country j will switch to R under the proposed alternative strategy. Thus, from all possibilities we conclude that country i is better off keeping P; this contradicts $\sigma(d_j, P) = R$, $\forall j \in \{1, 2, ..., N\}$. Thus, it must be true that $\sigma(d_j, P) = P$ $\forall j \in \{1, 2, ..., N\}$.

One implication of Proposition 3 is that it significantly narrows the set of possible pure strategy symmetric Markovian equilibria. For example, it rules out equilibria such as hegemons that always keep the status quo regime while dominants always switch (and also its mirrored version in which hegemons always switch while dominants keep the status quo). In fact, it reduces the set of possible pure strategy symmetric Markovian equilibria to only four possibilities: rules-based, power-based (which can be of two subclasses), and cyclic equilibria (see Definition 3).[19]

Corollary 1 (Rules, Power or Cyclic Equilibria) *Any symmetric Markovian equilibrium in pure strategies must be rules-based, power-based, or cyclic.*

The previous results can be illustrated with a numerical example. We simulate a two-country system that satisfies the assumptions on the static payoffs $u(\omega, \rho)$ described in equation (1). For simplicity, we consider a symmetric transition matrix Q.[20] Moreover, in line with the theoretical and empirical evidence discussed in section 2, we consider that static aggregate welfare (net of c) is higher under *rules* than under *power*.[21] We then check for which values of (δ, c) the system has a rules-based symmetric Markovian equilibrium. The

[19]There are $4^2 = 16$ possible pure strategy symmetric Markov strategy profiles, so our result reduces the set of possible classes of equilibria from 16 to 4.

[20]A symmetric transition matrix Q is such that $q_{h_i,d_i} = q_{h_j,d_j}, q_{d_i,h_i} = q_{d_j,h_j}$ and $q_{d_i,d_j} = q_{d_k,d_l}$, $\forall i, j, k, l \in \mathcal{I}$, with $i \neq j \neq k \neq l$.

[21]Specifically, the transition probabilities used for this numerical exercise are: $q_{h_1,h_1} = q_{d_1,d_1} = q_{h_2,h_2} = q_{d_2,d_2} = 0.1$, $q_{h_1,d_1} = q_{h_2,d_2} = 0.9$, $q_{d_1,h_1} = q_{d_1,d_2} = q_{d_2,h_2} = q_{d_2,d_1} = 0.45$, and $q_{h_1,d_2} = q_{h_1,h_2} = q_{d_1,h_2} = q_{d_2,h_1} = q_{h_2,h_1} = q_{h_2,d_1} = 0$. The resulting stationary distribution of this parametrization is $(\frac{1}{6}, \frac{1}{3}, \frac{1}{6}, \frac{1}{3})$, meaning that the long-run probability of being at a dominant state is $\frac{2}{3}$, while the probability of being at a hegemon state is $\frac{1}{3}$. For the static payoffs, we use: $u_i(h_i, R) = u_i(h_j, R) = u_i(d_i, R) = u_i(d_j, R) = 1$, $u_i(h_i, P) = u_i(h_j, P) = 0$, $u_i(d_i, P) = 1.1$, and $u_i(d_j, P) = 0.8$. Hence, global payoff is maximal under R, in which case the gains from trade are split evenly between the two countries; is minimal when a hegemon chooses P; and is intermediate when a dominant chooses P, in which case the dominant country obtains a larger share of the gains than the subordinate country.

result of this simulation is shown in Figure 6. In line with Corollary 1, the figure shows that all symmetric Markovian equilibria in pure strategies are rules-based, power-based or cyclic.

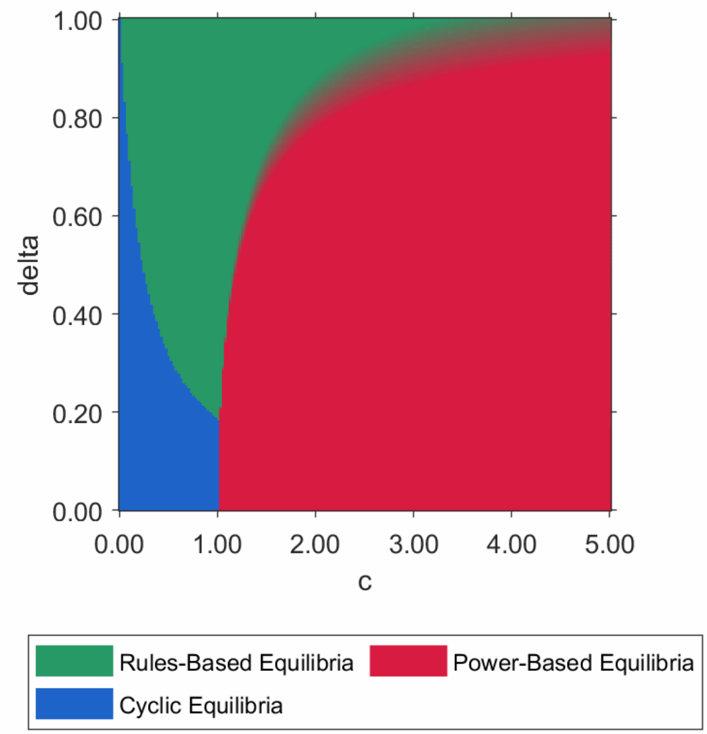

Figure 6: Symmetric Markovian equilibria

We confirm that, for sufficiently small c, a rules-based equilibrium does not exist (Proposition 2.A). Instead, there is cyclic equilibria (Proposition 1.A). For higher values of c in the frontier between cyclic and rules-based equilibria, there is a negative relationship between δ and c, illustrating that rules-based equilibria can only exist if c is not too small and that this minimum cost is a negative function of the discount factor – that is, with a higher δ, the cutoff $\underline{c}$ that makes a rules-based equilibrium unviable is lower (Proposition 1.A). The reason is that a greater value for future payoffs makes a dominant country more willing to keep R when it receives R.

For a sufficiently high c, a rules-based equilibrium does not exist (Proposition 2.B). Instead, there is power-based equilibria (Proposition 1.B). The kink at $c \approx 1$ illustrates the point at which it ceases to be myopically optimum for a hegemon to switch to R. At that point, a hegemon must exhibit sufficient foresight to be willing to transition to regime

R. That explains why the relationship between the cost $\bar{c}$ and the discount factor δ that defines the region where there is a rules-based equilibrium becomes positive after that point – that is, with a higher δ, the cutoff $\bar{c}$ that makes a rules-based equilibrium unviable is larger (Proposition 1.B). Furthermore, the figure shows that, if countries are very myopic (very low δ), then a rules-based equilibrium cannot exist (Proposition 1.C).

Interestingly, for a range of costs close to the threshold $\bar{c}$ that makes hegemonic countries indifferent, when the status quo is P, between keeping P and replacing it with R, a rules-based equilibrium exists only in mixed strategies. This corresponds to the shadowed region between the rules-based (green) and power-based (red) regions, where the colors are mapped into the strategies so that the "greener" is the area, the higher the probability that the hegemon switches to R if it receives P. A rules-based equilibrium in mixed strategies is one in which play at all states h_i, $\forall i$, involves mixing. In that equilibrium, any hegemon, upon receiving P as the status quo trade regime, randomizes between keeping P and switching to R. To see the intuition for a mixed strategy equilibrium, start with a set of parameters that yields a power-based equilibrium – in which case hegemons that receive P keep P. Then reduce c. Eventually, c will become small enough such that a hegemon that receives P will want to incur c to implement R. This may resemble a rules-based strategy. However, at the c that makes the current hegemon just indifferent between P and R upon receiving P, $\tilde{c}$, there is a discontinuity. This happens because we are focusing on symmetric equilibria, so that, for c just below $\tilde{c}$, if we consider that the current hegemon chooses R upon receiving P, we are also considering that future hegemons will behave in the same way – but that will affect the continuation payoff of the current hegemon. As a result, there is a range of c, given other parameters, where the only equilibria has every hegemon randomizing their regime choice upon receiving P.

A second implication of Proposition 3 is more fundamental. It shows that, regardless of any long-run strategic motive, and regardless of how small c may be, dominant countries will not introduce rules in equilibrium. Therefore, it implies a pivotal role for hegemons in creating a rules-based regime:

Corollary 2 (Indispensable Hegemon) *The world can switch from regime P to regime R only when it is in a hegemonic state.*

This result emphasizes the fundamental role played by hegemons in creating rules-based regimes. In that sense, Corollary 2 offers support for the main premise of Kindleberger's (1973) Hegemonic Stability Theory, along the lines argued by Brooks and Wohlforth (2016), that a hegemon is needed to establish a rules-based system. However, unlike what that theory prescribes, it does not imply that a hegemon is needed to *keep* a rules-based system in place.

Instead, the rules-based regime's long-term stability requires compliance from *both* hegemon and dominant countries. Dominant countries face immediate incentives to dismantle rules in favor of power-based systems, but these short-run incentives are countered by two long-term strategic forces. First, a dominant country must consider its potential future return to hegemony – abandoning rules today would force it to bear the costs of recreating it later. Second, the transient nature of leadership means that today's dominant power could soon find itself subordinate in a power-based system it helped create, leaving it worse off than under a rules-based order. These dynamics create conditions where, despite short-term advantages, dominant countries may choose to preserve rules for their long-term benefit.

To further investigate the dependence of the rules-based system on hegemon countries to create it, we show that our results extend to an environment in which there is a small exogenous probability of a rules breakdown at every period. This could be interpreted as an unexpected one-off shock that depletes the value of previous investment in setting up the rules. For example, technological developments (digital trade, the rise of services trade, the growth of global value chains, etc.) could make the current rules obsolete.[22] Then, to have in place a set of effective rules, the leader would need to incur in c again.

Specifically, suppose that at the beginning of every period when $\rho = R$, before the leading country chooses the regime, there is a small probability $\varepsilon \to 0$ that the regime will exogenously revert to *power*. Equation (2) then becomes:

$$\begin{aligned} V_i(\omega,\rho) = \max_{\{P,R\}}\{&u_i(\omega,P) + \delta \sum_{\omega'\in\Omega} q_{\omega,\omega'} V_i(\omega',P), \\ &u_i(\omega,R) - \mathbb{1}_{\{\rho=P\}}c + \delta \sum_{\omega'\in\Omega} q_{\omega,\omega'}[(1-\varepsilon)V_i(\omega',R) + \varepsilon V_i(\omega',P)]\}. \end{aligned}$$

It can be shown that all previous results hold in this world with ε-shocks, including Proposition 3 and its corollaries.

This makes clear the fragility of a rules-based system: once undermined, it can be restored only by a hegemon, even if c is very low, even if countries are very forward-looking, and even if aggregate welfare is higher under *rules*.

3.3 Rare Hegemons

Although hegemons are indispensable for implementing a rules-based regime, the *frequency* in which hegemons arise is not as critical for its sustainability. On one hand, because hegemons benefit from *rules*, their rarity may make sustaining *rules* more difficult. On the other hand,

[22] Indeed, Mavroidis (2025) argues that the current WTO rules are outdated. This is also a central message of Mattoo et al. (2025), but due to geopolitical rivalry motives.

the frequency of leadership turnover is also a key force for maintaining *rules*; if hegemons are rare but leadership changes are frequent, a rules-based system may still be sustained. To examine these two forces separately, we compare two distinct cases in which hegemons are *rare* (i.e., when their frequency in the long-run is sufficiently close to zero).

3.3.1 Enduring Dominance

First, we study the case in which there is "enduring dominance," i.e., once a country becomes dominant, it remains dominant for an arbitrarily long time. In this case, hegemons are rare (the steady state probability of a hegemon is close to zero) and there is very little leadership turnover. As a result, as we shut down the two channels that favor the sustainability of rules-based regimes, in this case, rules-based equilibria do not exist.

Proposition 4 (Enduring Dominance) *For any given $\delta < 1$, there exists $\bar{q} < 1$ such that, if $q_{d_i,d_i} > \bar{q}$, for any i, then there is no symmetric Markovian equilibrium in which regime R is chosen on-equilibrium path by dominant countries. In particular, rules-based equilibria do not exist.*

To illustrate Proposition 4, we consider the same baseline parameters as in Figure 6, but increase the probability that a dominant leader remains dominant in the subsequent period. Specifically, we increase that probability sequentially from $q_{d_i,d_i} = 0.1, i = 1, 2$, as set in the baseline simulation, to 0.35, 0.6, 0.85, and 0.99. In each step, we reduce the probabilities $q_{d_i,d_j}, i \neq j$, and $q_{d_i,h_i}, i = 1, 2$, by half of the increment in q_{d_i,d_i}, so that the transition probabilities starting from any state remain equal to one. Figure 7 shows the simulations in each of those cases. Clearly, as the probability that a dominant country remains dominant increases, the set of parameters under which a rules-based equilibrium emerges decreases. In particular, the minimal δ that allows for a rules-based equilibrium rises. If that probability reaches 0.99, the possibility of such an equilibrium virtually vanishes, even for δ very close to one.

Hence, if hegemons are rare (in a probabilistic sense) and leadership turnover is low, it is impossible to sustain a rules-based system.

3.3.2 High Leadership Turnover

We now study the case of rare hegemons coupled with high leadership turnover. Our goal is to show that it is possible to have rules-based equilibria even in worlds in which hegemons are rare.

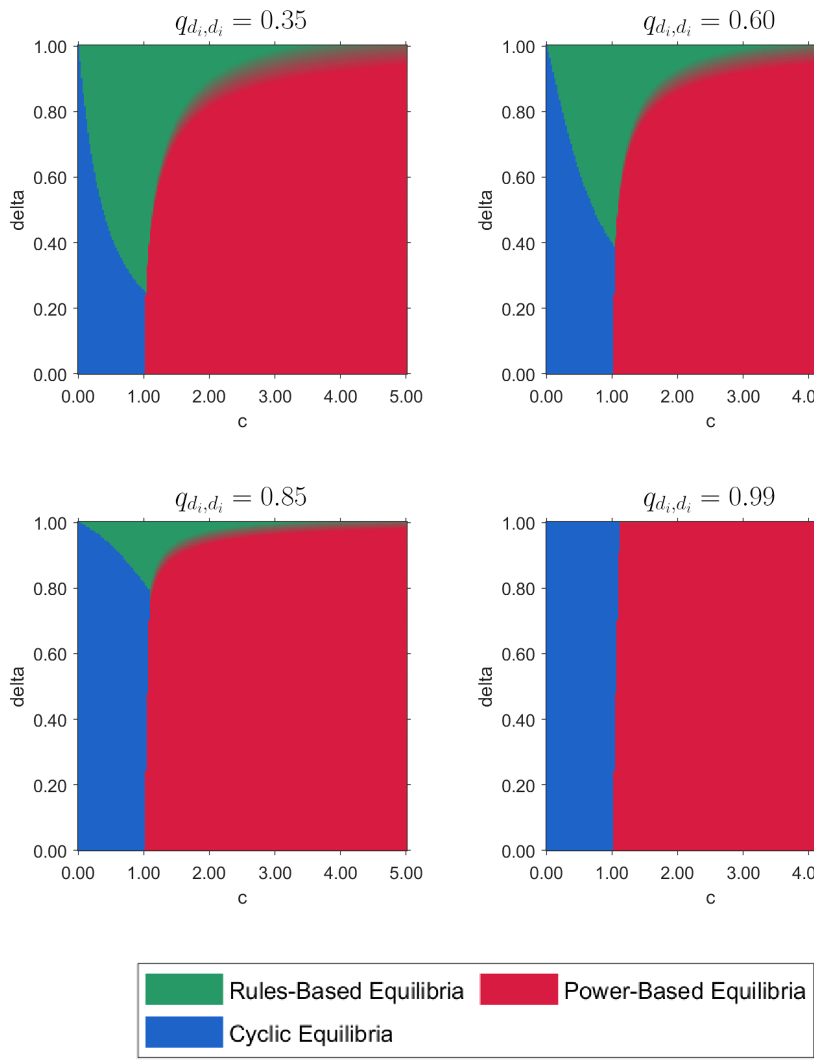

Figure 7: Enduring dominance

To do so, we focus on the two-country case and consider a special class of symmetric transition matrices,[23] which we refer to as high-leadership-turnover (HLT) matrices. These matrices feature very high transition probabilities out of dominant states into all other dominant states. Formally:

Definition 4 (High Leadership Turnover) *A transition matrix Q is an HLT matrix if*

$$\begin{cases} q_{d_i,d_i} = 0 & \forall i \\ q_{d_i,d_j} = 1 - \varepsilon & \forall i \neq j \\ q_{d_i,h_i} = \varepsilon & \forall i. \end{cases} \tag{HLT}$$

[23]See footnote 20 for the definition of a symmetric transition matrix Q.

We say that worlds with high leadership turnover have rare hegemons if the probability of a hegemon in the long-run is close to zero, where the steady state distribution over hegemon states is denoted by

$$\mu_H \equiv \sum_{i=1,2} \mu_{h_i}.$$

Under HLT, this steady state distribution μ_H is an increasing function of ε and hegemons are rare in the long-run if ε is small:[24]

$$\lim_{\varepsilon \to 0} \mu_H = 0.$$

We make the following two assumptions on payoffs. First, we assume that imposing *rules* is efficient for every state.

Assumption (Efficiency). For any $\omega \in \Omega$:

$$\sum_{i=1}^{N} u_i(\omega, R) > \sum_{i=1}^{N} u_i(\omega, P). \tag{EFF}$$

A consequence of (EFF) in our setting is that

$$\begin{aligned} 0 &< (u_i(d_i, R) - u_i(d_i, P)) + (u_j(d_i, R) - u_j(d_i, P)) \\ &= (u_i(d_i, R) - u_i(d_i, P)) + (u_i(d_j, R) - u_i(d_j, P)), \end{aligned}$$

where the equality follows from the symmetry of the payoffs. For ease of notation, let $\Delta u_i(\omega) := u_i(\omega, R) - u_i(\omega, P)$ and $\Delta V_i(\omega) := V_i(\omega, R) - V_i(\omega, P)$ for any $\omega \in \Omega$. The previous inequality can then be rewritten as:

$$\Delta u_i(d_i) + \Delta u_i(d_j) > 0.$$

Second, we assume that imposing *rules* under hegemony benefits the hegemon more than the average improvement of imposing *rules* under dominance between the leader and the subordinate.

Assumption (Hegemon benefits are large). For any i,

$$\Delta u_i(h_i) > \frac{\Delta u_i(d_i) + \Delta u_i(d_j)}{2} \tag{HBL}$$

[24]For the two-country case this steady state distribution is $\mu_H = \frac{\varepsilon}{1 - q_{h_i, h_i} + \varepsilon}$.

We now show that, in worlds with symmetric transition matrices Q that satisfy (HLT) and payoffs that satisfy (EFF) and (HBL), it is possible to have rules-based equilibria even when hegemons are rare.

Proposition 5 (Rules with Rare Hegemons and High Leadership Turnover) *Consider a transition matrix Q that satisfies (HLT) and suppose that assumptions (EFF) and (HBL) hold. Then, for ε small enough, there exists $\delta \in (0,1)$ and $c > 0$ such that a rules-based equilibrium exists.*

Proposition 5 makes clear that, even if hegemons are rare, a rules-based equilibrium can still exist, provided that leadership turnover is high enough. Thus, while hegemons are indispensable to create rules-based regimes, in line with the HST, their long-run sustainability does not hinge on the prevalence of a hegemon state, in contrast with the main thrust of the HST.

3.4 Efficiency

An important question is whether/when inefficiencies arise in the choice of international trade regimes. For example, since the only payoff restrictions are those in equation (1), our previous results show that a rules-based equilibrium may not exist even in a world in which aggregate welfare under a rules-based regime is higher than aggregate welfare under a power-based regime in *every* state of the world (i.e., when $\sum_{i\in\mathcal{I}} u_i(\omega, R) > \sum_{i\in\mathcal{I}} u_i(\omega, P)$, $\forall\omega \in \Omega$) and the cost c to introduce *rules* is relatively small. Would that be an inefficient outcome?

We formalize the problem of finding the efficient solution as follows. Consider a hypothetical "global social planner" with the authority to select the trade regime in every state, aiming to maximize the total payoff across all countries. The transition matrix Q induces a probability distribution over the sequence of possible states. The efficient solution is given by the choice of a trade regime at every period, possibly as a function of all previous states and choices. With slight abuse of notation, we say that the efficient solution is the sequence $\{\rho_t(\omega_t)\}_{t=1}^{\infty}$ that solves:

$$\max_{\{\rho_t(\omega_t)\}_{t=1}^{\infty}} \mathbb{E} \sum_{t=1}^{\infty} \delta^{t-1} \{ \sum_{i\in\mathcal{I}} u_i(\omega_t, \rho_t) - \mathbb{1}_{\{\rho_t = R \,\&\, \rho_{t-1}=P\}} c \}. \tag{SP}$$

As we have done in the numerical simulations, in this subsection we focus on the case where aggregate welfare in all states is higher under a rules-based regime than under a power-based regime, as this is the most natural and interesting case. Then two conclusions quickly emerge. First, the social planner will not replicate the cyclic equilibria, as these involve incurring the cost c repeatedly. Second, the social planner will either never pay the cost c or it will pay it once and keep the rules-based regime thereafter. Using these two insights, we can construct a function $\delta(c)$, increasing in c, such that for any pair (δ, c) with $\delta \geq \delta(c)$, the efficient solution is rules-based. In contrast, for any pair (δ, c) with $\delta < \delta(c)$, the efficient solution is power-based.

To formalize the efficient solution below, it is convenient to define the set $S \subset \mathbb{R}^2$ as the set of parameters (δ, c) such that the efficient solution induces *power* at every period and state.

Proposition 6 (Efficient Solution) *Assume that* $\sum_{i \in \mathcal{I}} u_i(\omega, R) > \sum_{i \in \mathcal{I}} u_i(\omega, P)$, $\forall \omega \in \Omega$. *The solution* $\{\rho_t(\omega_t)\}_{t=1}^{\infty}$ *of (SP) involves either the permanent choice of* P *or the permanent choice of* R. *Thus, the cost of creating rules is either paid once and rules are maintained thereafter, or it is never paid. Moreover,* S *is non-empty and satisfies the following properties:*

1. *If* $(\delta, c) \in S$, *then every point* (δ', c') *with* $\delta' \leq \delta$ *and* $c' \geq c$ *also belongs to* S.

2. *For any* $(\delta, c), (\delta', c') \in S$, $\lambda(\delta, c) + (1 - \lambda)(\delta', c')$, $\forall \lambda \in (0, 1)$, *also belong to* S.

Proposition 6 is illustrated by Figure 8, which replicates Figure 6 while adding a white curve that represents the function $\delta(c)$ that defines the contour of the set S. The social planner chooses *rules* in the region to the left of $\delta(c)$ and *power* to its right. When we compare the equilibria in the decentralized game with the efficient solution, there are multiple sources of inefficiencies. First, when countries are not very patient, hegemons avoid paying the cost of constructing a rules-based regime, knowing that future dominant countries will dismantle it. Second, when the cost c is very small, dominant countries are tempted to move to power-based regimes, knowing that hegemons will revert back to rules (the cyclic equilibria). Importantly, for intermediate values of c, power-based equilibria may arise despite the social planner's preference for a rules-based system. This discrepancy occurs because dominant countries prioritize their own interests, neglecting the welfare losses they impose on subordinate countries when maintaining *power*. Thus, we find that decisions regarding trade regimes are often inefficient and possess a systematic bias against sustaining a rules-based equilibrium, yielding power-based and cyclic equilibria in many circumstances where the efficient solution would be the permanent choice of rules.

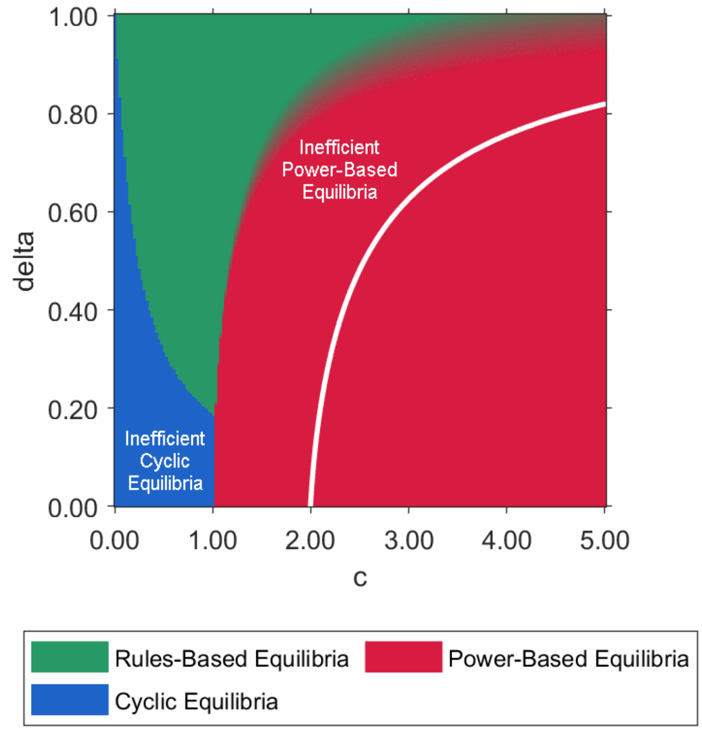

Figure 8: Efficient and inefficient equilibria

3.5 Maintaining a Rules-Based Regime

We have established that a hegemon is necessary to create a rules-based regime (Corollary 2) and that no hegemon will do so if the cost is too high (Proposition 2). Moreover, even if such a regime is efficient, no single country – hegemon or not – may be willing to bear the implementation costs. This raises a critical question: if a rules-based system were established due to an exceptional historical event – such as a rare moment of multilateral cooperation after a global war – would countries at least be willing to maintain it? If not, then even extraordinary circumstances would be insufficient to sustain a rules-based regime, making it fundamentally untenable. But if countries would uphold it once in place, then although countries may resist incurring the initial cost, the regime becomes self-enforcing. In that case, the sustainability of a rules-based system remains a viable prospect.

To investigate these questions, we distinguish the strategies that induce a power-based equilibrium into two different classes. First, a class of power-based strategies that is resistant to rules: a rules-based regime will not be created, and even if it were created, it would not persist. Second, a class of power-based strategies that do not induce the creation of a rules-based regime, but if it were created, then it would persist. The former strategy represents a world in which countries desire neither to create nor to maintain a rules-based system. The latter strategy represents a world where the cost of creating a rules-based regime is too large, but there is a desire to maintain it once it is created. We denote the first class of power-based strategies as *rules-resistant power-based* (RRP) and the second class as *rules-compatible power-based* (RCP).

Formally, these two types of strategies are distinguished by the prescribed play at off-equilibrium path histories. Intuitively, RRP describes an equilibrium that, following the off-equilibrium path play of R, reverts back to a power-based regime once a dominant country becomes the leader. Specifically:

Definition 5 (Rules-resistant Power-based) *An RRP strategy profile is defined as*

$$\sigma(\omega,\rho)=\begin{cases} R & \text{if } (\omega,\rho)=(h_i,R) \\ P & \text{if } (\omega,\rho)=(h_i,P) \\ P & \text{if } (\omega,\rho)=(d_i,\rho),\ \forall\rho. \end{cases} \tag{RRP}$$

Observe that this is consistent with the definition of power-based equilibrium in Definition 3, so it is a strategy that induces a power-based equilibrium.

In contrast, RCP describes an equilibrium where R is maintained endlessly following any off-equilibrium path history in which R is played. Specifically:

Definition 6 (Rules-compatible Power-based) *An RCP strategy profile is defined as*

$$\sigma(\omega,\rho)=\begin{cases} P & \text{if } (\omega,\rho)=(\omega,P),\ \forall\ \omega. \\ R & \text{if } (\omega,\rho)=(\omega,R),\ \forall\ \omega. \end{cases} \tag{RCP}$$

The crucial difference between strategies RRP and RCP is that they imply very different long-run responses to a potential play of R. Figure 9 distinguishes the cases in which the power-based equilibrium is either rules-resistant or rules-compatible. Note that a straight horizontal line separates RCP from RRP. This is because those equilibria differ on whether the dominant country is willing to sustain R or not upon receiving R, and thus the incentive constraints are independent of the cost c.

Note that RCP and RRP equilibria can be efficient or inefficient. Left of the $\delta(c)$ curve, and below the straight horizontal line, there is a region of inefficient power-based equilibria where a rules-based regime would not persist even if it were put in place exogenously.

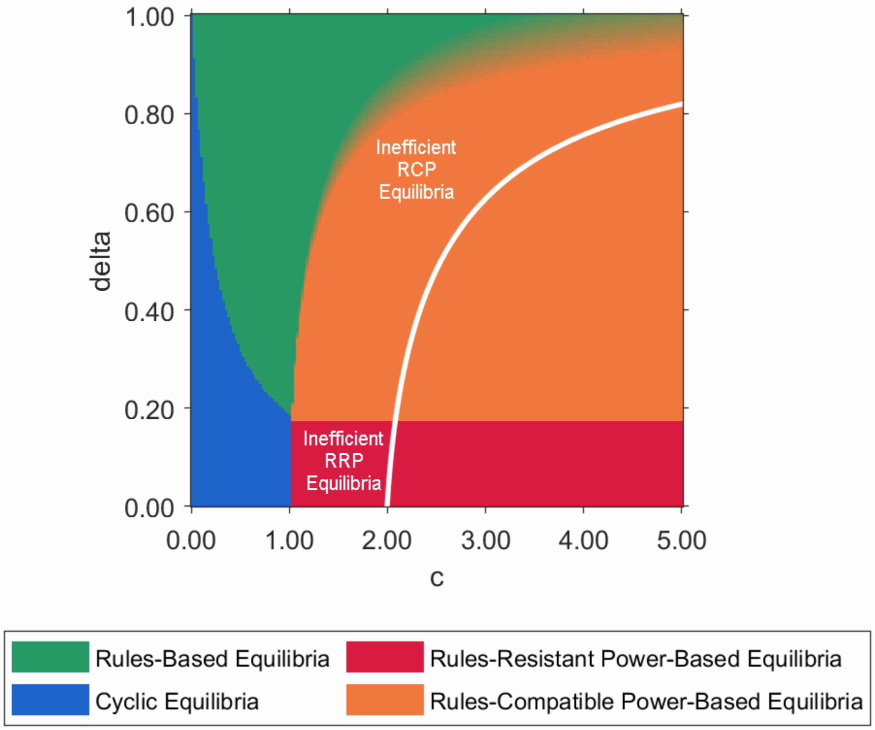

Figure 9: Rules-resistant and rules-compatible power-based equilibria

The more interesting region is the one where a rules-based regime is not an equilibrium but satisfies two key properties: (i) it is efficient, and (ii) once in place, it is self-enforcing (i.e., the region of inefficient RCP equilibria). This region is to the left of the $\delta(c)$ curve, but above the straight horizontal line. It reflects a collective action problem, as those properties indicate that transitioning from a power-based equilibrium to a rules-based system would be efficient, and once established, all countries would have a unilateral incentive to uphold it without requiring further negotiations. The challenge lies in the fact that no leading country is willing to bear the initial cost of establishing the rules-based system.

The distinction between RCP and RRP is also useful for analyzing one-off, unexpected shocks to governments' discount factor δ. Suppose the world is initially in a rules-based equilibrium and experiences a drop in δ, making governments less forward-looking. As illustrated in Figure 9, if the reduction in δ is "moderate," *rules* would still be upheld, since the cost c has already been incurred. Only when the decline in δ is sufficiently large does the system enter the RRP region, leading to a permanent shift toward *power*. This highlights that, once a rules-based equilibrium is established, it is resilient to unanticipated changes in governments' planning horizons—unless the shock is severe.

More broadly, the concepts of RCP and RRP equilibria also help clarify the role of the assumption that, in a static sense and net of c, hegemons prefer R to P (equation (1)). If we were to abandon this assumption, there would be no meaningful distinction between dominant and hegemon leaders. In light of our results, it would then follow that a rules-based regime would never emerge within our framework. We would instead need to posit that such a regime had been established under exceptional circumstances, as discussed at the beginning of this section. Even so, we know that in the region characterized by RCP, *rules* would still be maintained. The rationale for why a leader who benefits from *power* in the short run nonetheless chooses to uphold *rules* lies, once again, in the desire to protect itself once it is no longer in power. In fact, it can be shown that the set of parameters consistent with RCP when there is no hegemon state expands as the likelihood of leadership turnover increases.

4 Endogenous Transitions

In the benchmark model of section 3, we focused on the case where the transition matrix Q is exogenous and independent of countries' choices. It is plausible, however, that the transition out of leadership at a given period may depend on the trade regime chosen at that period, in which case leading countries would choose the trade regime taking that effect into account. Here we consider an extension of our model in which the permanence in a dominant position and the transition out of dominance (q_{d_i,d_i} and q_{d_i,d_j}, respectively) depend on the choice of the trade regime.

Specifically, we assume that when a dominant country chooses P, its probability of remaining dominant increases. That is, for any given initial trade regime ρ:

$$\tilde{q}_{d_i,d_i} = q_{d_i,d_i} + \gamma \quad \text{if } \sigma(d_i, \rho) = P,$$

for $\gamma \geq 0$.[25] For simplicity, we assume that q_{d_i,h_i}, $\forall i$, remains unchanged regardless of the choice of the regime, so that all counter-adjustments from the increase in q_{d_i,d_i} are in the transition from a dominant state to another dominant state. Our analysis so far has assumed $\gamma = 0$; we are now interested in the case where $\gamma > 0$. This is, originally, the most natural case to consider for endogenous transition probabilities. Under a power-based regime, larger

[25] We are focusing on departures from transition matrices that are independent of trade regime choices, so we consider $\gamma \in [0, \sum_{j \neq i} q_{d_i,d_j}]$, $\forall i$.

countries obtain (static) payoffs more aligned with their economic sizes. Arguably, this could endogenously prolong the power asymmetries, relative to the benchmark case in which transition probabilities are exogenous.

It is straightforward to show that our previous results remain valid in this modified environment. However, this change in the structure of the transition matrix alters the parameters under which each type of equilibrium emerges. The direct "endogenous dominance effect" is intuitive: when a dominant leader chooses P, it not only reaps the immediate benefits of a power-based regime, but also increases its likelihood of retaining dominance. As a result, P becomes relatively more attractive than R for dominant leaders compared to the benchmark model. The magnitude of this effect is increasing in γ.

However, there are two opposing indirect effects at work as well. First, as γ rises, the probability that a dominant leader keeps its leadership increases, which in turn raises the likelihood of eventually attaining hegemonic status. Since a hegemon is better off under R, this creates an incentive for the dominant leader to favor R over P. Second, if country i can prolong its dominance by choosing P, the same logic applies to country $j \neq i$. This implies that losing leadership when the prevailing regime is P becomes more costly, as the new dominant j will choose P if the status quo is P and by doing so j will also increase its likelihood of retaining leadership. Thus, while *power* helps sustain dominance, it also makes the prospect of losing leadership more detrimental, which discourages the dominant country from choosing *power*.

It then follows that, even if perhaps counterintuitively at first, allowing for $\gamma > 0$ does *not* necessarily make a rules-based equilibrium less likely. The net effect of increasing γ is, in general, ambiguous. For example, using the same baseline parametrization as Figure 6, we find that increasing γ has minimal impact on the equilibrium regions. Figure 10 illustrates this point most starkly by setting $\gamma = \bar{\gamma} \equiv 1 - \sum_{j \neq i} q_{d_i,d_j}$ under a symmetric Q. This implies that a dominant country choosing P entirely eliminates the risk of losing its leadership status – whether as dominant or hegemon. In this case, the second indirect effect vanishes, removing strategic interactions. Nevertheless, the conditions for a rules-based equilibrium remain largely unchanged compared to those in Figure 6. In fact, when c and δ are relatively large, a rules-based equilibrium is actually easier to sustain under $\gamma = \bar{\gamma}$ than under $\gamma = 0$. When countries value the future significantly and the cost of implementing R is high, a hegemon's willingness to incur c ensures that dominant countries prefer to maintain R rather than switching to P to secure indefinite leadership. This avoids a situation where, upon eventually becoming hegemon, they either face a low payoff under P or must pay c to reinstate R.

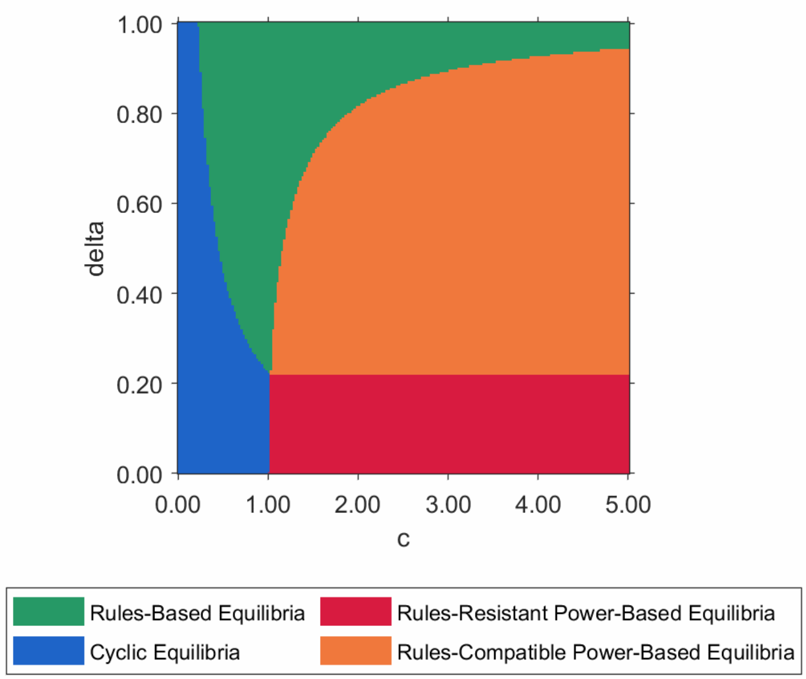

Figure 10: Equilibria under Endogenous Transitions, $\gamma = \bar{\gamma}$

In summary, while a power-based regime can reinforce dominance by increasing asymmetry between countries, it does not inherently threaten the sustainability of a rules-based equilibrium. The reason is that the direct benefits of choosing *power* are counterbalanced by indirect considerations, which often work to offset those gains.

5 A Bipolar World

We now extend the model to incorporate a state in which there are two leading countries instead of one. In this extension, a dominant country that loses its sole leadership moves into a state in which it shares leadership with another country: the equivalent of having simultaneously two dominant countries. We denote this extra state by Φ_{ij}. Thus, in a two-country version, we now have $\Omega = \{h_1, d_1, \Phi_{12}, d_2, h_2\}$. We maintain the assumption that the transitions are smooth, that is, all transitions to neighbor states are positive and $\forall i$, $q_{h_i,h_i} + q_{h_i,d_i} = 1$, $q_{d_i,h_i} + q_{d_i,d_i} + q_{d_i,\Phi_{ij}} = 1$ and $q_{\Phi_{ij},d_1} + q_{\Phi_{ij},\Phi_{ij}} + q_{\Phi_{ij},d_2} = 1$. Figure 11 illustrates this situation. In the general N-country case, there are also states $\Phi_{13}...\Phi_{1N}$, $\Phi_{23}...\Phi_{2N}$, $...\Phi_{N-1,N}$, and the smoothness assumption is adapted accordingly.

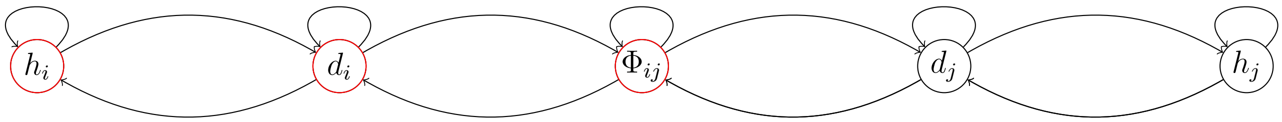

Figure 11: Transition between states; bipolar world

Whenever the system reaches state Φ_{ij}, the two dominant countries, i and j, play a simultaneous regime-choice game. In this game, each dominant simultaneously chooses whether to implement *rules* or *power*. As before, if the status quo is *rules*, there is no cost in maintaining it; otherwise, if a dominant country chooses to support *rules*, it will pay a cost. Specifically, when the status quo is P, if both countries coordinate on *rules*, their static payoff is $u_i(\Phi_{ij}, (R, R)) - \mathbf{1}_{\{\rho=P\}}c$. If they coordinate on *power*, they each get $u_i(\Phi_{ij}, (P, P))$, which is assumed to be smaller than $u_i(\Phi_{ij}, (R, R)) - c$. The matrix below represents these payoffs when the status quo regime is P, where i's regime choice is indicated by the first element of (ρ_i, ρ_j):

	R	P
R	$u_i(\Phi_{ij}, (R, R)) - c$, $u_j(\Phi_{ij}, (R, R)) - c$	$u_i(\Phi_{ij}, (R, P)) - c$, $u_j(\Phi_{ij}, (P, R))$
P	$u_i(\Phi_{ij}, (P, R))$, $u_j(\Phi_{ij}, (R, P)) - c$	$u_i(\Phi_{ij}, (P, P))$, $u_j(\Phi_{ij}, (P, P))$

Table 1: Bipolar state with status quo regime P

Three new issues arise in this context: (i) the need for coordination; (ii) exploitation of market power; and (iii) free-riding incentives. First, if *rules* requires both dominant countries to sustain it, then they would need to coordinate their efforts, which may be difficult.

Moreover, whenever a dominant country behaves cooperatively, as would be expected under *rules*, the other dominant country could gain in the short run by exploiting its power in international markets, designing its trade policies to improve its terms of trade at the expense of the other country. This force can be made explicit if we assume that, when the status quo is R, for all $i, j, i \neq j$,

$$u_i(\Phi_{ij}, (P, R)) > u_i(\Phi_{ij}, (R, R)) > u_i(\Phi_{ij}, (P, P)) > u_i(\Phi_{ij}, (R, P)).$$

These payoffs generate a prisoner's dilemma in the static game, in which each country has a dominant strategy – to choose power – but the outcome in which both choose *power* is Pareto dominated by the outcome in which they coordinate on *rules*. This represents the idea that a rules-based regime is jointly optimal for the two leaders in a static sense, but if one plays *rules* while the other dominant country does not, exploiting its market power, then the former would be worse off than it would have been had it abandoned *rules* itself.[26]

Finally, free-riding incentives emerge if one dominant country is able to sustain *rules* alone; this could compel the other dominant country to free ride on the first's institutional efforts.

We analyze each of these issues in turn. First, to consider the difficulties imposed by strategic coordination, we assume that:

Assumption A1. Maintaining or establishing a rules-based regime requires the joint support of both leaders.

This shuts down free-riding incentives with respect to the establishment/maintenance of *rules*. In this context, we find that there can be a rules-based equilibrium in pure strategies. Therefore, the need for strategic coordination does not eliminate the possibility of an enduring rules-based regime.

Proposition 7 (Rules-Based Equilibrium When *Rules* Requires Joint Support) *Under assumption A1, there can be a rules-based equilibrium in pure strategies.*

Under A1, the cost of deviating from a strategy that prescribes cooperation in maintaining *rules* in this state with two leaders is high: it implies a transition to a power-based system. This high cost of deviation enables, under some parameters, an equilibrium in which both leaders sustain *rules* in this state.

Now, to study the impact of market power, we maintain A1 and ask how an increase in the short-run gain from choosing *power* at the bipolar state affects the likelihood of a rules-based equilibrium.

Proposition 8 (Market Power Hinders Rules-Based Equilibria in Bipolar World) *Assume A1 and suppose that there is a rules-based equilibrium. In particular, at state Φ_{ij}, both leaders choose R when the status quo is R. Then decrease $u_i(\Phi_{ij}, (P, R))$. A rules-based equilibrium also exists under this new structure of payoffs. Moreover, for sufficiently high $u_i(\Phi, (P, R))$, a rules-based equilibrium cannot exist.*

Proposition 8 shows that large short-run gains from exploiting market power in international markets in a bipolar state can be critical to undermining the viability of a rules-based equilibrium.

[26] This illustrates a recurrent complaint of the U.S. government about China's policies, particularly its opaque system of domestic subsidies. See Boullenois et al. (2025) for a detailed account of China's subsidy and other state-support policies.

Finally, to shed light on free-riding, we drop assumption A1 and consider, instead, that:

Assumption A2. Only one leader is needed to maintain or establish a rules-based regime. Thus, if the choice of the leaders were (P, R) or (R, P), next period's status quo will be R.

Assumption A2 may appear, at first, more conducive to *rules* than Assumption A1, since A1 has a stricter requirement for the maintenance of a rules-based regime than A2 (i.e., the need for joint support for *rules* instead of a single leader's support). However, precisely because under A2 a single leader is enough for maintaining *rules*, a strong negative result emerges: a pure strategy rules-based equilibrium cannot exist in a world where two leaders can co-exist, and one can sustain *rules* alone. The reason is that the temptation to free ride on the opponent's efforts at maintaining *rules* (or in creating them) inhibits the sustainability of *rules* in a symmetric equilibrium. If one dominant country plays R, the other faces no long-run penalty by playing P; therefore, any short-run gain, no matter how small, is enough to induce the play of P. Since this rationale applies to both leaders, it is impossible to sustain *rules* in a bipolar state precisely when only one of them is needed to sustain it.

Now, a mixed strategy equilibrium in which each dominant country chooses rules with a positive probability smaller than one may exist. Still, there is always a positive probability that, upon reaching state Φ_{ij}, a world under a rules-based regime will switch to a power-based system.

Proposition 9 (Free-riding in Bipolar State Begets Power) *Under A2, a pure-strategy rules-based equilibrium does not exist. Consequently, we must have* $Pr(\sigma(\Phi_{ij}, \rho) = (P, P)) > 0$, $\forall \rho \in \{P, R\}$ *in any symmetric Markovian Equilibrium.*

Thus, free-riding incentives undermine the viability of pure-strategy rules-based equilibria. A rules-based regime is still possible, but only in a probabilistic sense. That is, in any symmetric Markovian equilibrium, *rules* will be displaced with positive probability at the bipolar state. This makes clear that, in a world with a bipolar state, relying on only one leader to sustain *rules* is likely to undermine its viability.

Observe that this negative result holds even when we isolate free-riding forces by setting market-power incentives to a minimum. To see that, consider that the following holds for $\epsilon \to 0$:

$$u_i(\Phi_{ij}, (P, R)) = u_i(\Phi_{ij}, (R, R)) + \epsilon > u_i(\Phi_{ij}, (P, P)) = u_i(\Phi_{ij}, (R, P)) + \epsilon.$$

This preserves the prisoner's dilemma structure for the short-run payoffs, but sets the differences in payoffs to nearly zero. And yet Proposition 9 still holds, highlighting the fragility of rules-based equilibria when there is a bipolar state.

To conclude this section, we present a result that relates HST to the context of a world with a bipolar state. Proposition 9 shows that there is no equilibrium in pure strategies in which *rules* is chosen on-equilibrium path in all states (including the bipolar state). However, an equilibrium in which hegemon and dominant leaders choose *rules* on-equilibrium path can exist (but, of course, not in the bipolar state). We prove the following necessary condition for such an equilibrium.

Proposition 10 (Hegemonic Stability Theory for Bipolar Worlds) *In a bipolar world, under assumption A2 and under a symmetric Q, a pure strategy symmetric Markovian equilibrium where hegemon countries create and maintain R and dominant countries maintain R exists only if the probability that a dominant country becomes hegemon is sufficiently high.*

Proposition 10 shows that, in a world in which there is a bipolar state and assumption A2 holds, dominant countries will sustain R only if the probability of moving to the hegemon state is sufficiently high. Intuitively, this happens because the only benefit that a dominant country has from retaining R is when it becomes a hegemon, since while dominant the country earns a lower payoff from R than from P, and if it loses its leadership, it cannot expect to secure R as the status quo (as we know from Proposition 9 that R will not be maintained in the bipolar state in a pure-strategy equilibrium).

6 A Multipolar World

In the benchmark model, we focus on a world in which there is always a leading country responsible for the current trade regime. In the previous section, we considered a world with the possibility of two concomitant leaders. We now include the possibility of a "multipolar" scenario, representing a state in which the countries defining the regime have little/no market power. We study the feasibility of a rules-based trade regime in such a world.

Specifically, suppose there is a state in Ω where neither a dominant nor a hegemon country exists. Denote this extended state space by $\tilde{\Omega} = \Omega \cup \{\mathcal{N}\}$, where $\mathcal{N}$ represents the extra state where no country is either hegemon or dominant. In this state, the choice of rules- versus power-based regimes depends on the actions of a group of n countries, where $n \in [2, N]$ is potentially a large number. For simplicity, let the identity of these n joint leaders be decided exogenously. This state is reachable after a period in which there is a dominant country, but not after a period in which there is a hegemon country. Specifically, $q_{h_i,\mathcal{N}} = q_{\mathcal{N},h_i} = 0$, $\forall i$, but $q_{d_j,\mathcal{N}} = q_{\mathcal{N},d_j} \geq 0$ for at least some j and $q_{\mathcal{N},\mathcal{N}} > 0$. Figure 12 illustrates this situation.

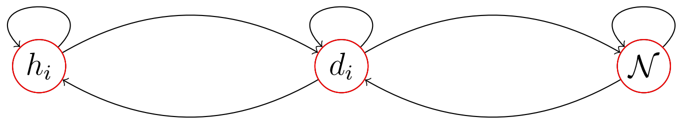

Figure 12: Transition between states; multipolar world

Whenever the system reaches state $\mathcal{N}$, all n joint leaders play a simultaneous regime-choice game. This game has the following properties. We assume that unanimity within this group is needed to have a rules-based regime at state $\mathcal{N}$. Let us also assume that the cost of supporting the switch from P to R is the same for all countries, given by $c' \leq c$.

A two-country version of this game, with country i being the row player and country j the column player, is shown below when the inherited regime is P:

	R	P
R	$u_i(\mathcal{N}, R) - c'$, $u_j(\mathcal{N}, R) - c'$	$u_i(\mathcal{N}, P) - c'$, $u_j(\mathcal{N}, P)$
P	$u_i(\mathcal{N}, P)$, $u_j(\mathcal{N}, P) - c'$	$u_i(\mathcal{N}, P)$, $u_j(\mathcal{N}, P)$

Table 2: Multipolar state with regime P

Each cell displays countries' static payoffs in each scenario. When all countries support R, each obtains the static payoff corresponding to R while paying cost c'. If at least one country does not support the switch, each country choosing P obtains the static payoff corresponding to P and each country that supports R also obtains the payoff corresponding to P while still paying c' for supporting the switch.[27] The analogous game when the inherited regime is R is

	R	P
R	$u_i(\mathcal{N}, R)$, $u_j(\mathcal{N}, R)$	$u_i(\mathcal{N}, P)$, $u_j(\mathcal{N}, P)$
P	$u_i(\mathcal{N}, P)$, $u_j(\mathcal{N}, P)$	$u_i(\mathcal{N}, P)$, $u_j(\mathcal{N}, P)$

Table 3: Multipolar state with regime R

The difference is that, when the status quo is R, countries do not need to incur c' to support keeping regime R.

We consider that $u_i(\mathcal{N}, R) > u_i(\mathcal{N}, P)$ $\forall i$, that is, in a multipolar state, c' aside, all countries prefer to be in a rules-based regime. Although the matrices above present only

[27] While the exact assumptions of how *rules* might be sustained and implemented in the multipolar state are somewhat ad hoc, the qualitative message of this section would remain valid if one relaxed some of the assumptions. For example, it may be the case that if P is implemented, then no country pays a cost, even if it had chosen R. We could also relax the unanimity assumption within the group of joint leaders and consider the case in which R is implemented in state $\mathcal{N}$ if a subset of countries strictly smaller than n chooses R.

the static payoffs, when considering the equilibrium continuation values discussed in the previous sections, at each combination of actions, the mutual choice of P is consistent with subgame perfect Nash equilibria. The reason is that, when all countries choose P, each gets a payoff of $V_i(\mathcal{N}, P)$, whereas if all other countries choose P except for i, then i's payoff is either the same as $V_i(\mathcal{N}, P)$ or lower. Additionally, depending on the value of c' and the equilibrium continuation payoffs, a scenario in which all countries choose R is also consistent with subgame perfect Nash equilibria.[28]

For the class of symmetric Markovian equilibria, there are only four action profiles in state $\mathcal{N}$ that are consistent with symmetric Markovian equilibria:

(i) **[Rules breakdown]** $\sigma(\mathcal{N}, P) = P$ and $\sigma(\mathcal{N}, R) = P$;

(ii) **[Coordination on R]** $\sigma(\mathcal{N}, P) = R$ and $\sigma(\mathcal{N}, R) = R$;

(iii) **[Regime reversal]** $\sigma(\mathcal{N}, P) = R$ and $\sigma(\mathcal{N}, R) = P$;

(iv) **[Keeping the status quo]** $\sigma(\mathcal{N}, P) = P$ and $\sigma(\mathcal{N}, R) = R$.

Action profile (i) represents an equilibrium in which whenever the system moves to state $\mathcal{N}$, countries keep the regime P if it were the status quo and move from R to P if R were the status quo. This rationale represents a situation in which a rules-based regime requires a level of leadership and coordination lacking in a multipolar world. Under such a strategy profile, a dominant country that considers choosing R now faces a positive probability of a regime breakdown in the subsequent period; thus, it may be tempted to choose P instead and profit from a higher short-term payoff.

A similar intuition holds for (ii), since in this case the regime choice in state $\mathcal{N}$ is R irrespective of the dominant country's precedent choice. Thus, a dominant country that would maintain R to hand R to the next dominant country may be tempted to switch to P, knowing that at $\mathcal{N}$ the regime will become R anyway. If the strategy profile induces (iii), then the incentive of a dominant country to keep R is even lower.

Thus, the most favorable scenario for the existence of an equilibrium in which a dominant country chooses R is if the equilibrium strategy induces (iv). That is, countries keep the status quo when in a multipolar state. We formalize this result below.

Lemma 2 (Maintaining Status Quo at Multipolar State) *Consider a world that has a multipolar state. If there exists a rules-based symmetric Markovian equilibrium in which*

[28]This last case could be, for example, an economy with many countries, each expecting that with very high probability they will be subordinate in the subsequent period, so that R is very attractive in state $\mathcal{N}$, provided that dominant countries keep it in equilibrium.

the strategy specifies action profiles (i), (ii) or (iii) in state $\mathcal{N}$, then there also exists a symmetric Markovian equilibrium that specifies profile (iv). The converse is not true.

From a technical perspective, Lemma 2 shows that focusing on strategies that maintain the status quo in state $\mathcal{N}$ gives us the largest set of parameters that allow for rules-based equilibria. More broadly, Lemma 2 also provides a normative lesson. Perhaps surprisingly, the long-run sustainability of a rules-based system is *diminished* when, in a multipolar state, the system always reverts to *rules*. Instead, what is desirable from that perspective is that countries adopt strategies in which they do not change the status quo regime in the absence of a leading country; this makes it more likely that such a regime will persist in the long run.

This has an important implication: rules-based equilibria exist for a broader set of parameters in a world that may transition to a multipolar state than in a world that may transition to a bipolar state.

Proposition 11 (Multipolar Admits More Rules-Based Equilibria than Bipolar) *Consider a world with a multipolar state and another world with a bipolar state. Consider also that both worlds have the same transition matrix Q.*[29] *Consider also that the payoffs in the intermediate states are the same under rules or power, but are larger in the bipolar world for a country that chooses power when the opponent chooses rules. Then, if there are no rules-based equilibria when the world has a multipolar state, there are no rules-based equilibria either when the world has a bipolar state.*

We establish this by showing that when countries adhere to the status quo in the multipolar state, the incentives to deviate to *power* when *rules* are in place are weaker than in a bipolar system. We then use the insight from Lemma 2 to conclude the proof.

The key mechanism underlying Proposition 11 is that deviations from R are more attractive in a bipolar state than in a multipolar one. This is because dominant countries in a bipolar world possess substantial market power – and thus strong incentives to manipulate their terms of trade –unlike leaders in a fragmented, multipolar world. As a result, it is easier for many small countries without market power to coordinate on R than for a few powerful countries with market influence to do so. From a broader – and more speculative – perspective, this insight suggests that the persistent difficulties in reaching new agreements within the WTO since the conclusion of the Uruguay Round in 1994 may not stem from the large number of members, as is often asserted in policy discussions (see, for example, Peres, 2024).

[29] That is, consider two worlds with the same number of countries N and same transitions $q_{\omega,\omega'}$ with the only difference being that the extra intermediate state is Φ_{ij} for the bipolar case and $\mathcal{N}$ for the multipolar case.

7 Reforming the System

So far we have taken the existing set of "rules" – which we interpret as those established under the GATT/WTO framework – as given. This helped us identify when a rules-based equilibrium is sustainable, but offered no guidance on how to reform the system when it is not. In practice, as the global economy evolves, so too must the rules governing it. Reform thus requires alternative frameworks that can address modern challenges while remaining both politically and economically feasible. We can formalize this in our model.

We consider three broad types of reforms, all motivated by an initial situation in which a dominant country receives *rules* but wishes to change the regime to *power*. The challenge is to design reforms such that the dominant country (as well as any future dominant) finds it optimal to maintain *rules* rather than abandon them. Against this backdrop, we study: (i) increasing the efficiency of the rules-based system; (ii) excluding the current dominant country from the rules-based system; and (iii) reshaping the system into a modified rules-based regime that redistributes the gains between leading and non-leading countries.

The first type of reform is theoretically the simplest. Few would argue that the WTO, its accomplishments notwithstanding, has exhausted all the gains from trade cooperation (Nicita et al., 2018). Indeed, recent proposals (e.g., Grossman and Sykes, 2025; Mavroidis, 2025) highlight new frameworks for WTO rules and disciplines that would presumably increase the efficiency of the system. We show here that efficiency and sustainability do move together:

Proposition 12 (Efficiency and Sustainability Move Together) *If the payoff of all countries under R, $u_i(\omega, R), \forall i$, increases (resp. decreases), then a rules-based equilibrium becomes possible under a larger (resp. smaller) set of parameters.*

Proposition 12 establishes a simple but important link: reforms that improve payoffs for all countries under *rules* both raise global welfare and make the system easier to sustain. Such Pareto-improving reforms may be difficult to design and implement in practice, though.

A second type of reform presumes that the current dominant country is disruptive (e.g., because it has an idiosyncratically low discount factor) and seeks to dismantle the system. The proposal is then to exclude that country from the system, while a subgroup of countries upholds the rules-based regime. This could be thought of as a "WTO - 1" system.[30] It has

[30] We are not the first to raise this possibility. Indeed, Horn and Mavroidis (2025) compellingly argue that, against the backdrop of the U.S.'s protectionist rules-less turn, the best alternative available would be a WTO without the U.S., even considering that rejoining in the future would require the U.S. to undergo the standard accession procedures. Krueger (2025) also endorses this approach, referring to it as the MUTO, or the "Minus U.S. Trade Organization," framework. In the policy arena, Mark Carney, Canada's Prime Minister, has similarly raised the possibility of a rules-based system that excludes the U.S. (*The Economist*, October 9, 2025).

two immediate consequences. First, excluding a large country generally reduces welfare for all. But if $u_i(\omega, R)$ falls for all remaining countries, we know from Proposition 12 that such a reform would weaken sustainability. Second, we have shown in the previous section that sustaining *rules* is easier under a multipolar than under a bipolar state (and surely also than under a disruptive dominant), because coordination among countries is not, in itself, a fundamental obstacle. This highlights the trade-off behind *WTO - 1.* If the world becomes multipolar, sustainability could be enhanced by neutralizing the incentives to undermine the system of a dominant leader. On the other hand, by forgoing the gains from trade with the largest economy, it risks undermining the very system it seeks to preserve.

A third type of reform involves changing the distribution of gains. This is related to a recurring theme in reform proposals (e.g., Rodrik, 2024; Stiglitz and Rodrik, 2024), which is is that a more flexible system could improve long-term viability.

A key reason why a rules-based system may fail is because dominant countries obtain higher static payoffs from *power* than from *rules*. If the rules are redesigned to narrow this payoff gap, dominants may be more willing to support *rules.* Improving the short-run gains of dominant countries, however, would likely reduce the benefits received by subordinates. Although non-leaders do not directly determine the regime, dominant countries accept short-term losses under a rules-based system only because they anticipate future gains once they lose dominance. If these future gains shrink, the long-run appeal of *rules* for dominants is weakened. Thus, a priori it is not clear that this would increase the appeal of *rules*.

Ultimately, a successful reform should seek the optimal balance between appeasing a dominant that threatens to dismantle the system and the sustainability of such a system. As we show next, such an "optimal appeasement" indeed exists.

Consider, then, a redesigned system of rules, denoted by *rules-lite* (RL) and defined as follows. For any country i,

$$u_i(\omega, RL) = \begin{cases} u_i(\omega, R) + \gamma, & \omega \in \{h_i, d_i\}, \\ u_i(\omega, R) - \lambda\gamma, & \omega \in \{h_j, d_j\},\ j \neq i, \end{cases} \tag{4}$$

where $\gamma \in [0, \bar{\gamma}]$, with $\bar{\gamma} \equiv u_i(d_i, P) - u_i(d_i, R)$ and $\lambda \in \left[1, \bar{\lambda}\right]$, $\bar{\lambda} \equiv \frac{u_i(d_j,R)-u_i(d_j,P)}{u_i(d_i,P)-u_i(d_i,R)}$. The payoffs in (4) imply that *rules-lite* never gives the leader more than it would receive under *power*, nor gives subordinates less than they would receive under *power* (both possibilities would make little economic sense). Parameter λ governs the degree of inefficiency of the reform. If $\lambda = 1$, the reform simply redistributes surplus from subordinates to leaders. If $\lambda > 1$, the reform is also wasteful. Here we take λ as given.

Proposition 13 (Optimal Appeasement) *Consider a symmetric world and a symmetric Markovian equilibrium in which a dominant country chooses* power *upon receiving* rules *(i.e., a cyclic equilibrium or a rules-resistant power equilibrium). There exists a unique threshold* $\gamma^\star$ *such that* $V_i(d_i, RL) \geq V_i(d_i, P)$ *if and only if* $\gamma \geq \gamma^\star$*. Hence, for* $\gamma \geq \gamma^\star$*,* rules-lite *is sustained by all leaders in their dominant states.*

At the extreme, when $\gamma = \bar{\gamma}$ and $\lambda = \bar{\lambda}$, *rules-lite* provides leaders with a payoff as large as they would obtain under *power* and subordinates with as little as they get under *power*. Such a regime is always sustainable. But it is also hollow: it mimics *power* in dominant states and improves upon it only in hegemon states. Sustainability, in this case, comes at the cost of desirability. The optimal appeasement reform corresponds to choosing $\gamma = \gamma^\star$, ensuring that dominant leaders are willing to maintain *rules-lite*.

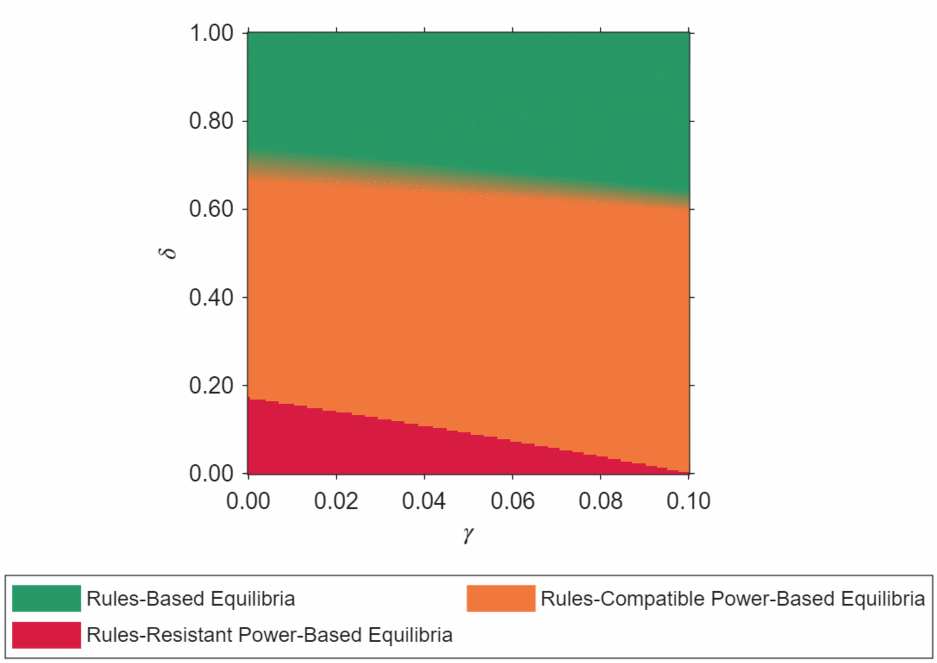

Figure 13: Sustainability of a Rules-Lite Regime with $c = 1.5$

To illustrate this result, we simulate our benchmark model using the parameterization from Figure 6 for a given c, but with a reallocation of payoffs as in (4), with $\lambda = \bar{\lambda} = 2$ and $\gamma \in [0, \bar{\gamma}]$, and in this case $\bar{\gamma} = 0.1$. At $\gamma = 0$, we recover the benchmark case. As γ rises, leaders' payoffs increase while subordinates' fall, making *rules-lite* inefficient relative to R.

At $\gamma = 0.1$, the system replicates the payoff under *power* in dominant states while improving on *power* in hegemon states. In this limiting case, rules-resistant equilibria disappear: *rules-lite*, once adopted, is always sustained. For intermediate γ, sustainability depends on the discount factor δ (Figure 13).

In sum, the three reform proposals face challenges of their own. Enhancing the gains from *rules* for all is clearly the most desirable, but it may be impracticable. Exclusion (*WTO - 1*) has the advantage of dethroning a disruptive leader, but reduces welfare for all and, because of that, could become unsustainable. Appeasement (rules-lite) can improve sustainability, but only by shifting surplus toward leaders and away from subordinates.

8 Conclusion

Over the last decade, the rules-based multilateral trading system has exhibited alarming signs of instability, as countries have increasingly disregarded established rules and leaned more heavily on economic power in managing trade relations. This trend intensified markedly during the second Trump administration. This is worrying, because that system is believed to have paved the way for significant gains from trade since the late 1940s. Furthermore, once dismantled, restoring the system may prove exceedingly difficult.

Arguably, this may be happening because the world no longer has a hegemonic power, as the U.S. used to be. To understand the risks to the rules-based trading system and how they relate to the absence of a hegemon, we use a dynamic stochastic game framework to model a situation where, in the presence of power asymmetry, the largest country has the prerogative to determine the trade regime, but its incentives depend on the extent of the asymmetry. We then analyze the equilibria in these long-term interactions and the viability of different trade regimes.

We find that the long-run sustainability of a rules-based trading regime is fragile. Once the world moves to a power-based regime, reverting to a rules-based system requires the presence of a hegemon. The maintenance of such a regime does not require a hegemon, but relies on enough turnover in world economic leadership. Our findings also underscore the stringent conditions necessary for the enduring viability of a rules-based system. The cost of establishing such a regime is a critical factor. If it is too small, regimes can shift easily, rendering the rules-based system unsustainable. Instead, the world would display a cyclic pattern, where rules would emerge only in the presence of a hegemon. Conversely, if the cost is too large, the world would be doomed to remain indefinitely in a power-based system.

When the world has two large countries with similar size, and each can affect the trade regime, additional challenges arise. In particular, we show that incentives to exploit market power in international markets and to free ride can be critical for the long-run sustainability of a rules-based regime. There may also be situations in which the world lacks a clear agenda-setter leader. We find that, because market power forces are weaker in this multipolar state, it is easier to sustain a rules-based equilibrium in that case than in an environment in which there is a bipolar state.

Overall, our findings underscore the risks posed by the ongoing erosion of multilateral cooperation and its potential long-term consequences. When viewed more broadly, our framework also helps explain recent shifts in multilateral trade cooperation. For decades after the 1940s, the U.S. was a clear hegemon, upholding a rules-based regime. However, its hegemony has steadily eroded over time. In the context of our model, the U.S. is now better characterized as a dominant leader with little prospect of regaining hegemony. At the same time, the rise of China has made a bipolar state increasingly likely. In such a state, market power and free-riding forces make it more difficult to sustain a rules-based regime. Anticipating this shift, the U.S. already faces incentives to move away from *rules*. Moreover, it is possible that we are not merely approaching a bipolar state but are already in one, hence the immense difficulties in sustaining the inherited rules-based system.

Is there hope? Our analysis clarifies what a successful reform of the rules-based system must accomplish to enhance its long-term viability. If a WTO reform were to benefit both leading and non-leading countries, it would generate the additional advantage of greater sustainability. If such an efficient reform proves infeasible, another possibility is to exclude the disruptive dominant, with a coalition of non-leading countries collectively managing the system. If this *WTO - 1* regime succeeds in preserving a rules-based regime, such a coalition would avoid the costly outcome of requiring a future hegemon to rebuild it from scratch. However, our analysis reveals that this is far from granted, because excluding a large country from the system would lower the gains from trade and thus also risk the sustainability of the system.

A third possibility highlights a fundamental trade-off in WTO reform: balancing concessions to leading countries against the costs borne by non-leading economies. In particular, one could design an "optimal appeasement" reform that grants the minimum necessary gains to leaders, while imposing the minimum possible losses on non-leading countries, subject to preserving the system's sustainability.

The central design problem is therefore to maximize global welfare while ensuring sustainability, recognizing that sustainability may come at the cost of a more unequal distribution of welfare gains and potential efficiency losses.

A The Hegemonic Stability Theory

A.1 The Hegemonic Stability Theory

The Hegemonic Stability Theory (HST) is an influential heuristic theory in the field of international relations and international political economy. It emerged during the 1970s as a response to the apparent connection between a stable global economic order and the presence of a hegemonic economic power. After World War II, the U.S. emerged as the hegemonic economic force. During that time, there was relative stability in the global economic system, with the Bretton Woods international monetary system and the GATT-based multilateral trade regime as the main legacies of this era.

However, in the 1970s, this stability seemed threatened by increasing economic conflicts among industrialized countries and the relative economic decline of the U.S. This environment drew attention to the idea that an open and stable global economy could be intrinsically linked to the existence of hegemonic power. It was in this context that the HST gained prominence through work by Charles P. Kindleberger, followed by contributions from Stephen D. Krasner and Robert O. Keohane, among others. It aimed to explain the inherent instability of non-hegemonic systems and the stability observed in hegemonic systems (Webb and Krasner, 1989).

The theory was initially proposed by Kindleberger (1973), who argues that the Great Depression of the 1930s resulted from a lack of leadership in providing global public goods such as discount facilities, countercyclical loans, or maintaining an open market for goods. During that period, as every country focused on protecting its own national interests, the global public interest deteriorated, creating an environment that rendered the system unstable.

It is in this context that Kindleberger (1973, p. 305) claims that "for the world economy to be stabilized, there has to be a stabilizer, one stabilizer." According to him, an open and stable world economy is causally related to the existence of a hegemon, which could provide certain institutional public goods that would benefit all countries regardless of whether they contributed to their production. Because the hegemon has a high enough stake at the benefits of this public good, it is willing to incur the cost to provide it.

Krasner (1976) extends and qualifies the original version of the theory, arguing that a hegemonic state is likely to lead to an open trading structure because it increases the hegemon's aggregate national income and its growth rate. He argues further that The likelihood of openness is highest during the ascendancy of a hegemonic state, as the decline of a hegemon brings with it a reduced ability to enforce compliance and a diminished interest in a system that could benefit others. In terms of our model, such a declining hegemon could be interpreted as what we define as a "dominant" country.

However, Krasner (1976) also notes that the costs and benefits of openness are not symmetric for all members of the system. While an open system may raise the absolute level of welfare of all participants, some countries will gain relative to others. If the distribution of relative gains is such that it poses a threat to the security of powerful nations, international liberalization will be restricted even if participation could enhance the absolute welfare of those nations.

Keohane (1982), as Kindleberger (1973), focuses on the provision of global public goods. He refers to this public good as a "regime," encompassing several international issue areas such as trade, monetary relations, energy, foreign investment, and security. According to him, countries respond rationally to constraints and incentives and, in the absence of authoritative global institutions, world politics may be characterized by an uncertain environment. For this reason, countries demand international regimes that aim to structure their relationships in stable and mutually beneficial ways.

Keohane (1984) agrees that hegemony plays a crucial role in establishing regimes, facilitating cooperation among countries. However, unlike Kindleberger (1973), he argues that hegemony is neither a necessary nor a sufficient condition for the emergence of cooperative relationships. Furthermore, once international regimes are established, cooperation, stability and openness do not necessarily require the existence of a hegemonic leader. His main hypothesis is that the demand for international regimes is in part a function of the effectiveness of these regimes, and the success of the institutions associated with a regime will itself become a source of regime persistence (Keohane, 1982). The evidence supporting this hypothesis lies in the fact that monetary and trade regimes endured even as America's hegemony waned in the 1970s, allowing for the potential of post-hegemonic cooperation.[31] While certain forms of cooperation tapered off, others persisted. This explains the lags between changes in power structures and changes in international regimes, and the reason why he emphasizes the greater ease of maintaining existing regimes than of creating new ones, as effective regimes can persist longer even in the absence of a leading power.

Like Keohane, Snidal (1985) argues that the decline or the absence of a hegemon does not necessarily lead to the collapse of economic order. Instead, it can create incentives for smaller states to engage in collective action, as they recognize the benefits of maintaining international institutions. Using a model of collective action, Snidal analyzes the theory's assumption that cooperation is impossible without a hegemonic power and finds that collective action is not only possible but may even increase as states work together to maintain

[31] It is important to highlight that, unlike Kindleberger's (1973) seminal contribution, Keohane's studies were developed during the 1980s, when there was already evidence that, even with the decline of U.S. hegemony, some successful regimes, such as the multilateral trade system based on rules, remained in force.

the order. Therefore, according to the author, hegemonic decline can actually strengthen cooperation by forcing states to collaborate.

In a more recent study, Brooks and Wohlforth (2016) examine the role of hegemonic leadership in promoting international order and cooperation, focusing on the United States' global role in the 21st century. According to them, earlier versions of the HST were often deterministic, leading some later theorists to downplay the importance of a hegemonic leader for maintaining and adapting international order and cooperation. Brooks and Wohlforth propose instead a probabilistic approach to hegemonic stability, arguing that an economically open, strongly institutionalized order is more likely to arise and be maintained when a hegemonic state actively pursues these objectives. They argue that these cooperative outcomes can occur without hegemonic leadership, but are significantly more likely to be achieved with it. This happens because the processes of order maintenance and creation are intrinsically connected. Once established, an institution must remain dynamic and capable of adapting to new challenges, and this is a task that often requires a hegemonic power to drive and support necessary revisions. Brooks and Wohlforth (2016) emphasize that this leader does not promote international order altruistically; it also extracts significant benefits for itself. Nevertheless, while the hegemon uses its power to primarily shape the system to serve its interests, this process also benefits other states.

A.2 HST and Equilibrium Trade Regimes

Our model assumes that there is power asymmetry in the international economic system; that the leading country has the prerogative to choose the trade regime; that there is a cost that the leading country pays to introduce a rules-based regime, which is higher than the cost to maintain it; and that the identity of the leading country may change over time. These are the main assumptions of the HST as well, although they are often left implicit. The key differences between our analysis and the HST are that (i) we make these assumptions explicit; (ii) we develop a formal dynamic model of forward-looking countries; and (iii) we characterize the Markovian equilibrium of a game among all countries. In our setting, the leading country is willing to sponsor a rules-based regime if the present value of the benefits is greater than the present value of the total cost *in an equilibrium of the dynamic game.* This is crucial because in a dynamic setting, the incentives of a forward-looking country regarding trade regimes can be very different from its static incentives, as in the HST.

In line with Kindleberger (1973) and Krasner (1976), we show that once the rules-based trade system is dismantled, the only way to establish it again is for the world to return to a hegemonic state. In that sense, we formalize that central prediction of the original version of the HST. But those authors also argue that an international environment without hegemonic power is synonymous of instability, as observed in the interwar period. We show that this is not necessarily true; it depends on the cost to (re-)establish a rules-based system, among other parameters. Instead, our results corroborate Keohane's (1984) view that, once international regimes are established, their maintenance becomes less costly, and need not require the existence of a hegemonic leader. Whether a rules-based system can be maintained without a hegemon depends on a set of parameters. We make explicit the conditions under which Keohane's departure from the original version of the HST holds, including in circumstances when there is no leading country, as in a multipolar world. Finally, as proposed by Brooks and Wohlforth (2016), our theoretical framework yields a probabilistic version of the HST.

B Other Figures

B.1 World Trade Policy Intervention Excluding the U.S. and China

The figure below illustrates the trend of global trade interventions from 2008 to 2023, excluding measures from the U.S. and China. The overall pattern is very similar to the one in the aggregate data (Figure 1), with protectionist measures consistently surpassing liberalizing ones, and with that difference increasing in the last years. This confirms that the shift toward protectionism over the past 15 years is not solely driven by the policies of the U.S. and China, but reflects a broader global trend.

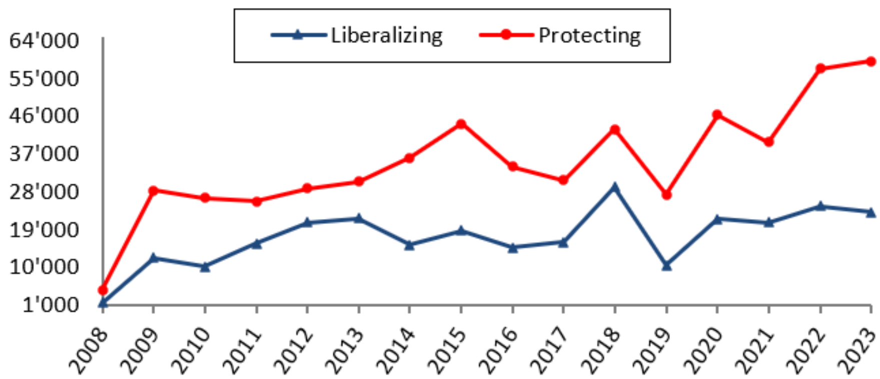

Figure 14: World trade policy interventions, excluding U.S. and China (2008–2023)
Source: Global Trade Alert database, 2024

C Britain Hegemony and Free Trade

By 1815, following the end of the Napoleonic Wars, Britain had emerged as the world's leading power, ushering in the era of Pax Britannica – a century of relative peace and stability in Europe and the world. One of the main features of Britain's hegemonic period was its leadership in trade liberalization, which reshaped the global economy.

Although Britain did not establish a formal rules-based trade system, it used its dominant position to influence global trade. Its shift away from mercantilism toward freer trade was gradual, beginning with the unilateral reduction of tariffs in response to evolving domestic and international pressures. The repeal of the Corn Laws in 1846 marked a key turning point in this transition (O'Brien and Pigman, 1992).

Britain's early success under liberalization enhanced its international standing, paving the way for its promotion of free trade ideology across Europe as a path to economic prosperity, peace, and stability (Kindleberger, 1975). In the 25 years following the repeal, countries such as France, Germany, Italy, Belgium, the Netherlands and Sweden also reduced trade barriers, presumably influenced by Britain's example. By the late 1850s, trade between Britain and France, Germany, and the U.S. had increased five- to sevenfold (Pigman, 1992).

After 1860, Britain adopted a more proactive trade policy strategy, moving from unilateralism to negotiated treaties based on mutual tariff reductions and the most-favored-nation clause (O'Brien and Pigman, 1992). In particular, the 1860 Cobden-Chevalier Treaty between Britain and France marked a turning point in Britain's trade policy. Its MFN clause ensured that trade concessions would be extended to all treaty partners. By 1865, Britain, France, Prussia and Italy had signed similar MFN treaties, which soon spread to smaller European nations (Kindleberger, 1975). This multilateral framework boosted market integration.

Between 1846 and 1870, Britain therefore used its hegemonic power, based on economic strength, military capacity, and global prestige, to lead global trade liberalization. According to the Hegemonic Stability Theory, a key public good provided by the hegemon is an open, free trading system, which in turn enhances the hegemon's own economic growth and political power (Krasner, 1976). Britain fulfilled this role by promoting liberal trade policies and negotiating MFN-based treaties that encouraged others to adopt similar principles.

After 1870, Britain's hegemonic position weakened as European and American industrial rivals emerged. Economic downturns, such as the Agricultural Depression, and conflicts like the Franco-Prussian War, fueled rising protectionism both abroad and at home. By the late 19th century, Britain shifted its focus toward defending imperial interests, and debates over free trade and tariff reform intensified. As trade wars and imperial rivalries escalated, liberalism gave way to conflict (Pigman, 1997). World War I and the interwar crises shattered Britain's economic leadership. By the 1930s, global trade had collapsed into protectionism, marking the end of an era shaped by British-led economic openness.

D Static Payoff Ordering

Figure 15 illustrates our rationalization for the static payoff ordering in a hegemonic state. It shows the utility frontier under a trade agreement when there are two countries and one – whose payoff is indicated in the Y-axis – is hegemonic.[32] If the hegemon chooses a power-based regime, because of the asymmetry of power, the gains from trade (GFT) would be skewed: closer to the Y-axis than to the X-axis, as indicated by point P_0. However, this ignores the possibility of renegotiation after irreversible investments have been carried out, when the hegemon would be able to extract more than the full GFT by threatening to dismantle the agreement, leaving the subordinate country worse than it would have been, had it not entered the trade agreement in the first place (point P_1). Anticipating this, the subordinate country would avoid such an agreement, and there would be no GFT to be shared. Faced with this participation constraint, the hegemon's best alternative is then to propose a rules-based system that neutralizes the effect of power asymmetry in negotiations and split the GFT more evenly. This outcome is represented by point R.[33] For the hegemon, the payoff obtained at point R is smaller than the payoff obtained at point P_0 (or P_1), but the latter is unachievable.

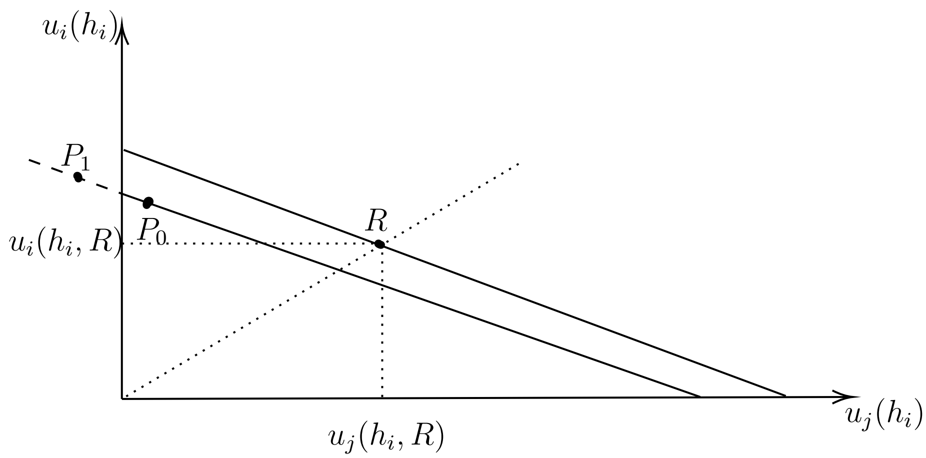

Figure 15: Stage game payoffs - Hegemon case

Figure 16 illustrates our rationalization for the static payoff ordering in a dominant state. It shows the utility frontier under a trade agreement when there are two countries and one – whose payoff is indicated in the Y-axis – is dominant. Now, after renegotiation the

[32]The linearity of the utility frontier reflects transferability of utility between countries, as would be the case with the availability of international transfers, for example. But this is not essential; it is assumed only for clarity.

[33]Observe that the figure depicts the GFT under *rules* as higher than under *power*, following the findings from the literature. However, we do not impose this restriction in most of our formal analysis.

subordinate country is still better off than without any trade agreement. Therefore, since the dominant country does not need to support a rules-based system to induce participation, it is better off by engaging in power-based negotiations, without offering the possibility of a rules-based regime to the subordinate country.[34]

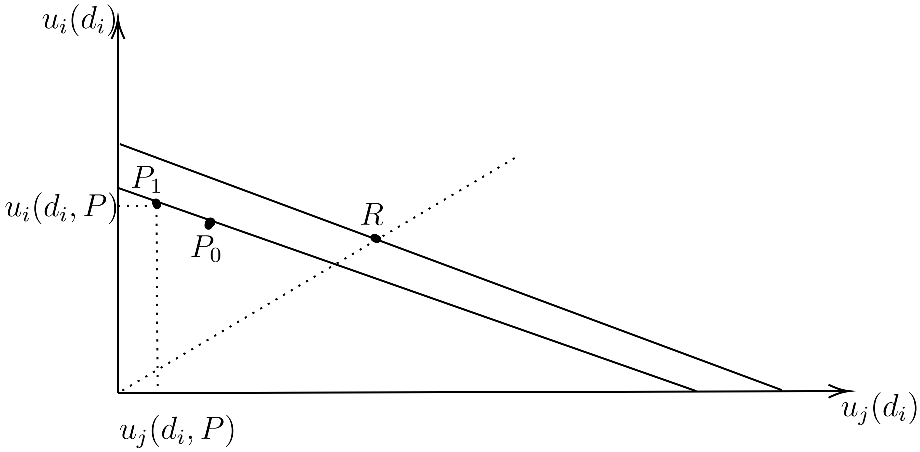

Figure 16: Stage game payoffs - Dominant case

[34]Observe that, as drawn, $u_i(d_i, P)$ in Figure 16 may look higher than $u_i(h_i, R)$ in Figure 15. But even if $u_i(d_i, P) > u_i(h_i, R)$, a hegemon may still not want to give up its hegemony to achieve a higher GFT. Here we are focusing on the gains from trade, and abstracting from many other advantages that a hegemon may enjoy that are not directly related to trade – from a higher standard of living regardless of trade connections to military ascendancy over the world, to other benefits of non-economic nature.

E Proofs

Proof of Proposition 1. For part A, consider a symmetric strategy profile $\sigma(\omega,\rho)$ in which $\sigma(h,\rho)=R$, and $\sigma(d,\rho)=P$, $\forall\rho$. Such a strategy generates a continuation payoff for every state (ω,ρ), denoted by $\tilde{V}(\omega,\rho)$. Before we proceed, note that this strategy σ induces $\tilde{V}_i(h_i,P)=\tilde{V}_i(h_i,R)-c$, $\tilde{V}_i(d_i,P)=\tilde{V}_i(d_i,R)$, $\tilde{V}_i(h_j,P)=\tilde{V}_i(h_j,R)$, $\tilde{V}_i(d_j,P)=\tilde{V}_i(d_j,R)$, $\forall i,j\in\{1,2,...,N\}$, with $i\neq j$. Therefore, succinctly, we have that $\tilde{V}_i(\omega,R)\geq\tilde{V}_i(\omega,P)$, $\forall i\in\{1,2,...,N\},\forall\omega\in\Omega$. Thus, $\forall c<u_i(h_i,R)-u_i(h_i,P)$ we can write:

$$u_i(h_i,R)+\delta\sum_{\omega'}q_{\omega,\omega'}\tilde{V}(\omega',R)-c\geq u_i(h_i,P)+\delta\sum_{\omega'}q_{\omega,\omega'}\tilde{V}(\omega',P). \tag{5}$$

In addition, note that

$$\begin{aligned}
&u_i(d_i,P)+\delta\sum_{\omega'}q_{d_i,\omega'}\tilde{V}(\omega',P)=\\
&u_i(d_i,P)+\delta\{q_{d_i,h_i}\tilde{V}(h_i,P)+\sum_{\omega'\in\Omega,\omega'\neq h_i}q_{d_i,\omega'}\tilde{V}(\omega',R)\}=\\
&u_i(d_i,P)-u_i(d_i,R)+u_i(d_i,R)+\delta\{q_{d_i,h_i}\tilde{V}(h_i,R)-q_{d_i,h_i}c+\sum_{\omega'\in\Omega,\omega'\neq h_i}q_{d_i,\omega'}\tilde{V}(\omega',R)\}=\\
&u_i(d_i,R)+\delta\sum_{\omega'}q_{d_i,\omega'}\tilde{V}(\omega',R)+(u_i(d_i,P)-u_i(d_i,R))-\delta q_{d_i,h_i}c.
\end{aligned}$$

Thus, for

$$\forall c<\frac{u_i(d_i,P)-u_i(d_i,R)}{\delta q_{d_i,h_i}}, \tag{6}$$

we have that:

$$u_i(d_i,P)+\delta\sum_{\omega'}q_{d_i,\omega'}\tilde{V}(\omega',P)\geq u_i(d_i,R)+\delta\sum_{\omega'}q_{\omega,\omega'}\tilde{V}(\omega',R). \tag{7}$$

Conditions (5) and (7) imply that the proposed strategy profile σ induces a symmetric Markovian equilibrium; indeed, it induces the cyclic equilibrium. Moreover, if we decrease δ, condition (6) still holds and so do conditions (5) and (7). We defer the proof of uniqueness to Lemma 3.

For part B, consider a strategy profile for which $\sigma(\omega,P)=P,\forall\omega\in\Omega$. Then, the induced payoffs are denoted by $\tilde{V}(\omega,\rho)$. This strategy will induce an equilibrium if:

$$u(\omega,P)+\delta\sum_{\omega'}q_{\omega,\omega'}\tilde{V}(\omega',P)>u(\omega,R)-c+\delta\sum_{\omega'}q_{\omega,\omega'}\tilde{V}(\omega',R),\ \ \forall\omega\in\Omega.$$

This can be simplified to:

$$c>u(\omega,R)-u(\omega,P)+\delta\sum_{\omega'}q_{\omega,\omega'}(\tilde{V}(\omega',R)-\tilde{V}(\omega',P)). \tag{8}$$

Moreover, since

$$\tilde{V}(\omega',R)-\tilde{V}(\omega',P)\leq\frac{\max_{\omega,\rho}u(\omega,\rho)-\min_{\omega,\rho}u(\omega,\rho)}{1-\delta},$$

we have that there exists $\bar{c}$ for which $\forall c>\bar{c}$, condition (8) is satisfied and thus there is a power-based equilibrium. We again defer the proof of uniqueness to Lemma 3. ■

Lemma 3 *Fix any given discount factor $\tilde{\delta}<1$, then, we have that:*

A. For c sufficiently small, the cyclic equilibrium is the unique symmetric Markovian equilibrium.

B. For c sufficiently large, the power-based equilibrium is the unique symmetric Markovian equilibrium.

Proof. There cannot be a power-based equilibrium since for c sufficiently small ($c<u_i(h_i,R)-u_i(h_i,P),\forall i$), the hegemon will always choose rules. To prove uniqueness, we rely on two arguments: (i) it cannot be a rules-based equilibrium (which we proved in Proposition 2); and (ii) if it is neither a rules-based nor a power-based, then it either does not exist or is a cyclic equilibrium (we prove this in Corollary 1). Part B is analogous to part A of this proof: for sufficiently large c, it cannot be the case that $\sigma(h_i,P)=R$ and, by Corollary 1, we are left with power-based as the only class of possible symmetric Markovian equilibria. ■

Proof of Proposition 2. Part A. Suppose, by contradiction, that a rules-based symmetric Markovian equilibrium exists $\forall$ c. Define $\Delta V_i(\omega)\equiv V_i(\omega,R)-V_i(\omega,P)$. Consider a cost $\tilde{c}$ such that $\tilde{c}<min_j\{u_j(h_j,R)-u_j(h_j,P)\}$. If $c<\tilde{c}$, an hegemon will choose regime R (more profitable in both the short and the long runs (Lemma 1)). In a rules-based equilibrium,

$$u_i(d_i,R)+\delta\sum_{\omega'\in\Omega}q_{d_i,\omega'}V_i(\omega',R)\geq u_i(d_i,P)+\delta\sum_{\omega'\in\Omega}q_{d_i,\omega'}V_i(\omega',P), \tag{9}$$

that is, $V_i(d_i, R) \geq V_i(d_i, P)$. We also know from Proposition 3 that if σ is a symmetric Markovian equilibrium, it must be such that $\sigma(d_i, P) = P$. Additionally, from the proof of Proposition 3 we know that in equilibrium we must have

$$u_i(d_i, P) + \delta \sum_{\omega' \in \Omega} q_{d_i,\omega'} V_i(\omega', P) > u_i(d_i, R) - c + \delta \sum_{\omega' \in \Omega} q_{d_i,\omega'} V_i(\omega', R), \tag{10}$$

that is, $V_i(d_i, P) > V_i(d_i, R) - c$. Combining (9) and (10) gives us the following inequalities:

$$c > \Delta V_i(d_i) \geq 0, \forall i. \tag{11}$$

For $c = 0$, inequality (11) cannot hold. Since the value functions V are continuous in c and $V_i(d_i, R)$ is not affected by c (in a rules-based equilibrium, once the system is in R, there is no other history in which rules will have to be constructed again on-equilibrium path) whereas $V_i(d_i, P)$ is decreasing in c, we have that for small c (in particular smaller than $\tilde{c}$) a symmetric Markovian rules-based equilibrium cannot exist.

Part B. A rules-based equilibria specifies that $\sigma(h_i, P) = R$. Thus, it requires

$$\begin{aligned} u_i(h_i, R) - c + \delta \sum_{\omega' \in \Omega} q_{h_i,\omega'} V_i(\omega', R) &\geq u_i(h_i, P) + \delta \sum_{\omega' \in \Omega} q_{h_i,\omega'} V_i(\omega', P) \\ \Leftrightarrow c \leq u_i(h_i, R) - u_i(h_i, P) &+ \delta \sum_{\omega' \in \Omega} q_{h_i,\omega'} \{V_i(\omega', R) - V_i(\omega', P)\}. \end{aligned}$$

We know that $u_i(h_i, R) - u_i(h_i, P) > 0$ and bounded, and also that $\forall \omega'$, $V_i(\omega', R) - V_i(\omega', P) > 0$ and bounded. Thus, there exists $\bar{c} > u_i(h_i, R) - u_i(h_i, P)$ for which $\forall c > \bar{c}$, the inequality above is violated.

Part C. We can write the rules-based equilibrium payoff at a dominant state j as:

$$\begin{aligned} V_i(d_i, R) &= u_i(d_i, R) + \delta \sum_{\omega} q_{d_i,\omega} V_i(\omega, R) \\ &\leq u_i(d_i, R) + \frac{\delta}{1-\delta} \bar{u}, \end{aligned} \tag{12}$$

where $\bar{u}$ is the maximum payoff a country can get in any state under any strategy. Moreover, to have a rules-based equilibrium, it must be the case that

$$\begin{aligned} V_i(d_i, R) &\geq u_i(d_i, P) + \delta \sum_{\omega} q_{d_i,\omega} V_j(\omega, P) \\ &\geq u_i(d_i, P) + \frac{\delta}{1-\delta} \underline{u}, \end{aligned} \tag{13}$$

where $\underline{u}$ is the minimum stage game payoff in any state and under any strategy. Note that (i) $\bar{u}$ and $\underline{u}$ are not functions of δ and (ii) $\Delta \equiv \frac{\delta}{1-\delta}(\bar{u}-\underline{u})$ is increasing in δ. Since $u_i(d_i, P) > u_i(d_i, R)$, there exists $\underline{\delta} > 0$ such that

$$u_i(d_i, P) > u_i(d_i, R) + \frac{\delta}{1-\delta}(\bar{u}-\underline{u}), \ \forall \delta < \underline{\delta}. \tag{14}$$

Therefore, for such $\delta < \underline{\delta}$, (14) contradicts (12) and (13). ■

Proof of Lemma 1. In states $\omega \in \{h_i, d_i\}$, upon observing $\rho = R$, country i can ensure itself of at least $V_i(\omega, R)$, but can also change to P at no cost and thus guarantee itself $V_i(\omega, P)$. Moreover, in any symmetric Markovian equilibrium, it must be the case that if $\sigma(\omega, R) = P$, then $\sigma(\omega, P) = P$. This can be seen directly from the inequalities below, representing country j's expected continuation payoffs. That is, if j is the leader at ω and $\sigma(\omega, R) = P$, then

$$u_j(\omega, P) + \delta \sum_{\omega' \in \Omega} q_{\omega,\omega'} V_j(\omega', P) \geq u_j(\omega, R) + \delta \sum_{\omega' \in \Omega} q_{\omega,\omega'} V_j(\omega', R).$$

It follows directly that

$$u_j(\omega, P) + \delta \sum_{\omega' \in \Omega} q_{\omega,\omega'} V_j(\omega', P) > u_j(\omega, R) - c + \delta \sum_{\omega' \in \Omega} q_{\omega,\omega'} V_j(\omega', R).$$

That is, if a country prefers not to choose R when it is costless, then it will also not choose it when it has a cost $c > 0$. Thus, in the states where i is subordinate, $\Omega \setminus \{h_i, d_i\}$, it is better for country i that the leader inherits regime R than regime P. If upon receiving regime R the leader prefers to set $\rho = P$, it will do so regardless of the regime it inherits. On the other hand, upon receiving P the leader may not want to switch from P to R because of the cost c. And, by assumption, the subordinate country is better off in the static game under $\rho = R$. ■

Proof of Proposition 3. The first part of the proposition follows from Lemma 1 and the fact that $u_i(h_i, R) > u_i(h_i, P)$, $\forall i$ (by assumption). Together, they imply that

$$u_i(h_i, R) + \delta \sum_{\omega' \in \Omega} q_{h_i,\omega'} V(\omega', R) \geq u_i(h_i, P) + \delta \sum_{\omega' \in \Omega} q_{h_i,\omega'} V(\omega', P), \tag{15}$$

so country i keeps R when it inherits R in state h_i.

To prove the second part of the result, let us suppose by contradiction that σ^* is a symmetric equilibrium strategy profile with $\sigma(d_i, P) = R$ for some $i \in \{1, 2, ..., N\}$. Since σ^* is a symmetric equilibrium, it must also be that $\sigma(d_j, P) = R$, $\forall j \in \{1, 2, ..., N\}$. We will construct a deviation strategy σ' and show that it improves upon the proposed strategy σ^*, implying that σ^* is not an equilibrium, a contradiction.

Let us construct this deviation strategy σ' in a way that it is identical to σ^* at all histories except (at most) two: (i) at a history h' in which the pair (ω, ρ) is (d_i, P) and (ii) at a history $h'' = (h', (h_i, P))$, which is defined to be a history at state h_i that follows history h' in conjunction with the play of P. At history h', this deviation strategy specifies the play of P (recall that the candidate equilibrium strategy σ^* specifies the play of R at such history), whereas at history h'' the constructed deviation strategy σ' specifies the play of R (which may or may not coincide with the action specified by σ^* at this history h'').

Consider country i's payoff having reached history h' and following the equilibrium strategy σ^*. This payoff is given by

$$u_i(d_i, R) - c + \delta \sum_{\omega' \in \Omega} q_{d_i,\omega'} V^{\sigma^*|h'}(\omega', R).$$

In contrast, the payoff following history h' under the deviation strategy σ' is given by

$$u_i(d_i, P) + \delta \sum_{\omega' \in \Omega} q_{d_i,\omega'} V^{\sigma'|h'}(\omega', P).$$

We will compare these two expected continuation payoffs and show that the payoff following the deviation strategy σ' is better than the payoff following the proposed equilibrium strategy σ^*.

We already know that $u_i(d_i, R) < u_i(d_i, P)$. Moreover, following this history h', which is at state d_i, we know that the subsequent state will be h_i, d_i, or d_j, for some $j \in \{1, 2, ..., N\}$ Let us look at each of these cases. If from state d_i, the system moves to state h_i, following the deviation strategy, the history will be h'', and the deviation strategy specifies the choice of R (and would incur a cost c). If, instead, the country had followed σ^*, it would have chosen R at the pair (d_i, P) and would have arrived at state h_i under regime R, and we know that $\sigma^*(h_i, R) = R$. And since we are focusing on Markovian equilibria, it then must be true that $V_{\sigma'|h''}(h_i, P)$ is the same as $V_{\sigma^*|h''}(h_i, R)$ except for the cost incurred, which under σ^* the cost is paid immediately, whereas under σ' the cost is paid in the subsequent period and $-c < \delta(-c)$. If the system stays at state d_i instead, which happens with probability q_{d_i,d_i}, the deviation strategy returns to the play specified by the original equilibrium strategy $\sigma'(d_i, P) = R$. Moreover, if $\sigma^*(d_i, P) = R$ it must be that $\sigma^*(d_i, R) = R$. Thus, an argument analogous to the one above holds, that is, $-c < \delta(-c)$. Finally, consider the case in which the system moves to state d_j. The deviation strategy σ' at this history is identical to σ^* and it specifies that country j will choose R, i.e. $\sigma^*(d_j, P) = R$. Moreover, under the equilibrium strategy σ^*, we would have arrived in state d_j under regime R, but $\sigma^*(d_j, R) = R$. Therefore, $V_{\sigma'}(d_j, P) = V_{\sigma^*}(d_j, R)$. Thus, we conclude that the deviation strategy σ' implies a higher current payoff for country i at history h' than the proposed equilibrium strategy σ^* and an identical continuation payoff. Thus, σ^* cannot be a subgame perfect Nash equilibrium. This is a contradiction. ■

Proof of Corollary 1. Given Proposition 3, we know that $\forall i$ it must be the case that $\sigma(h_i, R) = R$ and $\sigma(d_i, P) = P$. Thus, to fully specify only a symmetric Markovian equilibrium remains to specify $\sigma(h_i, P)$ and $\sigma(d_i, R)$. There are four possibilities: (i) $\sigma(h_i, P) = P$ and $\sigma(d_i, R) = P$; (ii) $\sigma(h_i, P) = P$ and $\sigma(d_i, R) = R$; (iii) $\sigma(h_i, P) = R$ and $\sigma(d_i, R) = R$; and (iv) $\sigma(h_i, P) = R$ and $\sigma(d_i, R) = P$. Cases (i) and (ii) are power-based, case (iii) is rules-based, and case (iv) is a cyclic equilibrium. ■

Proof of Proposition 4. The first useful result is to note that for any given δ and any given equilibrium, there exists Δv such that $V(\omega, R) - V(\omega, P) \leq \frac{\Delta v}{1-\delta}, \forall \omega$. Moreover, recall that there exists $\underline{u}$ such that $V(\omega, P) \geq \frac{\underline{u}}{1-\delta}, \forall \omega$, and $\bar{u}$ such that $V(\omega, R) \leq \frac{\bar{u}}{1-\delta}, \forall \omega$.[35] Then, for any equilibrium, we can write the expected continuation payoff at (d_i, P) as:

$$V_i(d_i, P) = u_i(d_i, P) + \delta q_{d_i,d_i} V_i(d_i, P) + \delta \sum_{\omega' \neq d_i} q_{d_i,\omega'} V_i(\omega', P)$$

$$V_i(d_i, P) = \frac{1}{1 - \delta q_{d_i,d_i}} \left\{ u_i(d_i, P) + \delta \sum_{\omega' \neq d_i} q_{d_i,\omega'} V_i(\omega', P) \right\}$$

$$V_i(d_i, P) \geq \frac{1}{1 - \delta q_{d_i,d_i}} \left\{ u_i(d_i, P) + \delta (1 - q_{d_i,d_i}) \frac{\underline{u}}{1 - \delta} \right\}$$

Moreover, the expected continuation payoff at (d_i, R) in any equilibrium in which $\sigma(d_i, R) = R$ can be written as:

$$V_i(d_i, R) = u_i(d_i, R) + \delta q_{d_i,d_i} V_i(d_i, R) + \delta \sum_{\omega' \neq d_i} q_{d_i,\omega'} V_i(\omega', R)$$

$$V_i(d_i, R) = \frac{1}{1 - \delta q_{d_i,d_i}} \left\{ u_i(d_i, R) + \delta \sum_{\omega' \neq d_i} q_{d_i,\omega'} V_i(\omega', R) \right\}$$

$$V_i(d_i, R) \leq \frac{1}{1 - \delta q_{d_i,d_i}} \left\{ u_i(d_i, R) + \delta (1 - q_{d_i,d_i}) \frac{\bar{u}}{1 - \delta} \right\}$$

[35]For example, consider $\Delta v \equiv \max_{\omega,\rho} u(\omega, \rho) - \min_{\omega,\rho} u(\omega, \rho)$, and $\bar{u} \equiv \max_{\omega,\rho} u(\omega, \rho)$.

Together, these two conditions imply that at any equilibrium in which rules is played on-equilibrium path, at state (d_i, R) we must have that:

$$u_i(d_i,R)+\delta q_{d_i,d_i}V_i(d_i,R)+\delta\sum_{\omega'\neq d_i}q_{d_i,\omega'}V_i(\omega',R)\geq u_i(d_i,P)+\delta q_{d_i,d_i}V_i(d_i,P)+\delta\sum_{\omega'\neq d_i}q_{d_i,\omega'}V_i(\omega',P)$$
$$u_i(d_i,P)-u_i(d_i,R)\leq\delta q_{d_i,d_i}\{V_i(d_i,R)-V_i(d_i,P)\}+\delta\sum_{\omega'\neq d_i}q_{d_i,\omega'}\{V_i(\omega',R)-V_i(\omega',P)\}$$

Which in turn implies that it is necessary that:

$$u_i(d_i,P)-u_i(d_i,R)\leq\delta\frac{q_{d_i,d_i}}{1-\delta q_{d_i,d_i}}\{u_i(d_i,R)-u_i(d_i,P)+\delta(1-q_{d_i,d_i})\frac{\bar{u}-\underline{u}}{1-\delta}\}+\delta(1-q_{d_i,d_i})\frac{\Delta v}{1-\delta} \tag{16}$$

As $q_{d_i,d_i}\to 1$, we can see that the r.h.s. of (16) converges to $\delta\frac{1}{1-\delta}\{u_i(d_i,R)-u_i(d_i,P)\}<0$. Since $u_i(d_i,P)-u_i(d_i,R)>0$, we have finalized the proof. ■

Proof of Proposition 5. A Rules-based equilibrium exists if and only if three conditions are satisfied: 1. A hegemon country must prefer to create Rules:

$$u_i\left(h_i,R\right)-c+\delta\sum_{\omega'}q_{h_i,\omega'}V_i\left(\omega',R\right)\geqslant u_i\left(h_i,P\right)+\delta\sum_{\omega'}q_{h_i,\omega'}V_i\left(\omega',P\right)$$

2. A dominant country prefers to maintain P when it receives P . (We know from Proposition 3 that this is a necessary condition for rules-based equilibria):

$$u_i\left(d_i,P\right)+\delta\sum_{\omega'}q_{d_i,\omega'}V_i\left(\omega',P\right)\geqslant u_i\left(d_i,R\right)-c+\delta\sum_{\omega'}q_{d_i,\omega'}V_i\left(\omega',R\right)$$

3. A dominant country prefers to maintain R rather than switching to P :

$$u_i\left(d_i,R\right)+\delta\sum_{\omega'}q_{h_i,\omega'}V_i\left(\omega',R\right)\geqslant u_i\left(d_i,P\right)+\delta\sum_{\omega'}q_{h_i,\omega'}V_i\left(\omega',P\right)$$

The three conditions above can be rewritten as:

$$c \leqslant \Delta u_i(h_i) + \delta \sum_{\omega'} q_{h_i,\omega'} \Delta V_i(\omega') \tag{1}$$

$$c \geqslant \Delta u_i(d_i) + \delta \sum_{\omega'} q_{d_i,\omega'} \Delta V_i(\omega') \tag{2}$$

$$0 \leqslant \Delta u_i(d_i) + \delta \sum_{\omega'} q_{d_i,\omega'} \Delta V_i(\omega') \tag{3}$$

In a rules-based equilibrium, a hegemon country creates rules, thus, it must be that: $\Delta V_i(h_i) = c$ and $\Delta V_i(h_j) = 0$, for all $i \neq j$. Therefore, for two countries, we can write, using (HLT):

$$\begin{aligned} \Delta V_i(d_i) &= \Delta u_i(d_i) + \delta \{\varepsilon \Delta V_i(h_i) + (1-\varepsilon)\Delta V_i(d_j)\} \\ &= \Delta u_i(d_i) + \delta \{\varepsilon c + (1-\varepsilon)\Delta V_i(d_j)\} \end{aligned}$$

Moreover,

$$\begin{aligned} \Delta V_i(d_j) &= \Delta u_i(d_j) + \delta \{\varepsilon \Delta V_i(h_j) + (1-\varepsilon)\Delta V_i(d_i)\} \\ &= \Delta u_i(d_j) + \delta \{(1-\varepsilon)\Delta V_i(d_i)\} \end{aligned}$$

This defines a linear system of two equations and two unknowns. Solving this system we have:

$$\begin{aligned} \Delta V_i(d_i) &= \frac{\Delta u_i(d_i) + \delta \varepsilon c + \delta(1-\varepsilon)\Delta u_i(d_j)}{1-\delta^2(1-\varepsilon)^2} \\ \Delta V_i(d_j) &= \frac{\Delta u_i(d_j) + \delta(1-\varepsilon)\left(\Delta u_i(d_i) + \delta \varepsilon c\right)}{1-\delta^2(1-\varepsilon)^2} \end{aligned}$$

Condition (2) can be rewritten as:

$$\begin{aligned} c &\geqslant \Delta u_i(d_i) + \delta \{\varepsilon \Delta V_i(h_i) + (1-\varepsilon)\Delta V_i(d_j)\} \\ &= \Delta u_i(d_i) + \delta \{\varepsilon c + (1-\varepsilon)\Delta V_i(d_j)\} \\ &= \Delta u_i(d_i) + \delta \varepsilon c + \delta(1-\varepsilon)\frac{\Delta u_i(d_j) + \delta(1-\varepsilon)\left(\Delta u_i(d_i) + \delta \varepsilon c\right)}{1-\delta^2(1-\varepsilon)^2} =: RHS_1(\delta, \varepsilon) \end{aligned}$$

However, as $\varepsilon \to 0$:

$$\begin{aligned} \lim_{\varepsilon \to 0} RHS_1(\delta, \varepsilon) &= \Delta u_i(d_i) + \delta \frac{\Delta u_i(d_j) + \delta \Delta u_i(d_i)}{1-\delta^2} \\ &= \frac{\Delta u_i(d_i) + \delta \Delta u_i(d_j)}{1-\delta^2} \end{aligned}$$

Note that, from (EFF), we know that $\Delta u_i(d_i) + \Delta u_i(d_j) > 0$, thus, for δ close enough to 1 , we have that $\lim_{\varepsilon \to 0} RHS_1(\delta, \varepsilon) > 0$.

Condition (1) can be rewritten as:

$$
\begin{aligned}
c &\leqslant \Delta u_i\left(h_i\right)+\delta\left\{q_{h_i,h_i}\Delta V_i\left(h_i\right)+\left(1-q_{h_i,h_i}\right)\Delta V_i\left(d_i\right)\right\} \\
&= \Delta u_i\left(h_i\right)+\delta\left\{q_{h_i,h_i}c+\left(1-q_{h_i,h_i}\right)\left[\frac{\Delta u_i\left(d_i\right)+\delta\varepsilon c+\delta(1-\varepsilon)\Delta u_i\left(d_j\right)}{1-\delta^2(1-\varepsilon)^2}\right]\right\} \Leftrightarrow \\
c\left(1-\delta q_{h_i,h_i}-\frac{\delta^2\varepsilon\left(1-q_{h_i,h_i}\right)}{1-\delta^2(1-\varepsilon)^2}\right) &\leqslant \Delta u_i\left(h_i\right)+\delta\left(1-q_{h_i,h_i}\right)\frac{\Delta u_i\left(d_i\right)+\delta(1-\varepsilon)\Delta u_i\left(d_j\right)}{1-\delta^2(1-\varepsilon)^2} \Leftrightarrow \\
c &\leqslant \frac{\Delta u_i\left(h_i\right)+\delta\left(1-q_{h_i,h_i}\right)\frac{\Delta u_i(d_i)+\delta(1-\varepsilon)\Delta u_i(d_j)}{1-\delta^2(1-\varepsilon)^2}}{1-\delta q_{h_i,h_i}-\frac{\delta^2\varepsilon\left(1-q_{h_i,h_i}\right)}{1-\delta^2(1-\varepsilon)^2}} =: \mathrm{RHS}_2(\delta,\varepsilon)
\end{aligned}
$$

However, as $\varepsilon\rightarrow 0$:

$$
\begin{aligned}
\lim_{\varepsilon\rightarrow 0}\mathrm{RHS}_2(\delta,\varepsilon) &= \frac{\Delta u_i\left(h_i\right)+\delta\left(1-q_{h_i,h_i}\right)\frac{\Delta u_i(d_i)+\delta\Delta u_i(d_j)}{1-\delta^2}}{1-\delta q_{h_i,h_i}} \\
&= \frac{\Delta u_i\left(h_i\right)+(\delta-1)\frac{\Delta u_i(d_i)+\delta\Delta u_i(d_j)}{1-\delta^2}}{1-\delta q_{h_i,h_i}}+\frac{\Delta u_i\left(d_i\right)+\delta\Delta u_i\left(d_j\right)}{1-\delta^2} \\
&= \frac{\Delta u_i\left(h_i\right)-\frac{\Delta u_i(d_i)+\delta\Delta u_i(d_j)}{1+\delta}}{1-\delta q_{h_i,h_i}}+\frac{\Delta u_i\left(d_i\right)+\delta\Delta u_i\left(d_j\right)}{1-\delta^2} \\
&> \frac{\frac{\Delta u_i(d_i)+\Delta u_i(d_j)}{2}-\frac{\Delta u_i(d_i)+\delta\Delta u_i(d_j)}{1+\delta}}{1-\delta q_{h_i,h_i}}+\lim_{\varepsilon\rightarrow 0}RHS_1(\delta,\varepsilon) \\
&= \frac{\Delta u_i\left(d_i\right)\left(\frac{1}{2}-\frac{1}{1+\delta}\right)+\Delta u_i\left(d_j\right)\left(\frac{1}{2}-\frac{\delta}{1+\delta}\right)}{1-\delta q_{h_i,h_i}}+\lim_{\varepsilon\rightarrow 0}RHS_1(\delta,\varepsilon) \\
&= \frac{\Delta u_i\left(d_i\right)\left(\frac{\delta-1}{2(1+\delta)}\right)+\Delta u_i\left(d_j\right)\left(\frac{1-\delta}{2(1+\delta)}\right)}{1-\delta q_{h_i,h_i}}+\lim_{\varepsilon\rightarrow 0}RHS_1(\delta,\varepsilon)
\end{aligned}
$$

Rare Hegemons

$$
\begin{aligned}
&= \frac{\left(\Delta u_i\left(d_j\right)-\Delta u_i\left(d_i\right)\right)(1-\delta)}{2(1+\delta)\left(1-\delta q_{h_i,h_i}\right)}+\lim_{\varepsilon\rightarrow 0}RHS_1(\delta,\varepsilon) \\
&> \lim_{\varepsilon\rightarrow 0}RHS_1(\delta,\varepsilon)
\end{aligned}
$$

for $\delta\in(0,1)$, where the first inequality follows from (HBL). Thus there exists ε small enough such that $RHS_2(\delta,\varepsilon)>RHS_1(\delta,\varepsilon)$. This means that there exists $c>0$ such that conditions (1) and (2) are satisfied. Finally, condition (3) can be rewritten as:

$$
\begin{aligned}
0 &\leqslant \Delta u_i\left(d_i\right)+\delta\left\{\varepsilon\Delta V_i\left(h_i\right)+(1-\varepsilon)\Delta V_i\left(d_j\right)\right\} \\
&= \Delta u_i\left(d_i\right)+\delta\left\{\varepsilon c+(1-\varepsilon)\frac{\Delta u_i\left(d_j\right)+\delta(1-\varepsilon)\left(\Delta u_i\left(d_i\right)+\delta\varepsilon c\right)}{1-\delta^2(1-\varepsilon)^2}\right\} = \mathrm{RHS}_3(\delta,\varepsilon)
\end{aligned}
$$

However, as $\varepsilon \to 0$:

$$\lim_{\varepsilon \to 0} RHS_3(\delta, \varepsilon) = \Delta u_i(d_i) + \delta \frac{\Delta u_i(d_j) + \delta \Delta u_i(d_i)}{1-\delta^2}$$
$$= \frac{\Delta u_i(d_i) + \delta \Delta u_i(d_j)}{1-\delta^2}$$

However, from (EFF), we know that $\Delta u_i(d_i) + \Delta u_i(d_j) > 0$, thus, for δ close enough to 1, we have that $\lim_{\varepsilon \to 0} RHS_3(\delta, \varepsilon) > 0$. Thus, there exists ε small enough such that condition (3) is satisfied. ■

Proof of Proposition 6. A switch from R to P at any moment in time clearly cannot be part of the efficient solution since we are assuming that $\sum_{i \in \mathcal{I}} u_i(\omega, R) > \sum_{i \in \mathcal{I}} u_i(\omega, P)$, $\forall \omega \in \Omega$. Thus, if the cost c of switching from P to R is ever paid, the efficient solution must be such that it will keep R from that moment onwards. Therefore, given any pair (δ, c) the efficient solution either prescribes a once-and-for-all switch from P to R or it prescribes the permanent play of P. Call the former case a payment case and the latter a no-payment case. Fix any pair $(\bar{\delta}, \bar{c})$. Suppose that it is a payment case. Then, for any (δ, c) with (i) $\delta = \bar{\delta}$ and $c < \bar{c}$, or (ii) $\delta > \bar{\delta}$ and $c = \bar{c}$, the efficient solution will also prescribe payment. If instead it prescribes no-payment, then for any (δ, c) with (i) $\delta = \bar{\delta}$ and $c > \bar{c}$, or (ii) $\delta < \bar{\delta}$ and $c = \bar{c}$, the efficient solution will also prescribe no-payment. To see this, suppose that at $(\bar{\delta}, \bar{c})$ the efficient solution prescribes payment. Then we can write:

$$\sum_{t \geq 1} \sum_{i \in \mathcal{I}} \delta^{t-1} u_i(\omega, R) \pi(\omega, t) - c \leq \sum_{t \geq 1} \sum_{i \in \mathcal{I}} \delta^{t-1} u_i(\omega, P) \pi(\omega, t),$$

where $\pi(\omega, t)$ denotes the probability of state ω at period t.

$$c \leq \sum_{t \geq 1} \sum_{i \in \mathcal{I}} \delta^{t-1} \{u_i(\omega, R) - u_i(\omega, P)\} \pi(\omega, t). \tag{17}$$

Given our assumption above, we have that $\sum_{i \in \mathcal{I}} u_i(\omega, R) - \sum_{i \in \mathcal{I}} u_i(\omega, P) > 0$ for all $\omega \in \Omega$. Thus, the right-hand side of the inequality (17) is decreasing in δ. Additionally, note that the set S of parameters (δ, c) such that the efficient solution prescribes the play of rules-based regime is non-empty. In particular, it includes $\forall c$ with $c < \sum_{j \in \mathcal{I}} \{u_j(h_i, R) - u_j(h_i, P)\}$, $\forall \delta$. It also includes pairs (δ, c) with high δ. Formally, $\forall c$, there exists δ_c such that $\forall \delta > \delta_c$, the efficient solution prescribes rules-based. This can be seen from the fact that the choice of rules yields an aggregate payoff that is higher than the aggregate payoff under power in every state, thus, as long as δ is sufficiently high, the difference between the discounted sum of payoffs under rules versus power will be higher than the cost c. Similarly, for any δ, a sufficiently high cost exists such that the efficient solution prescribes the constant choice of a power-based regime. Suppose not. Then, there exists i for which

$$\frac{\sum_{j\in\mathcal{I}} u_j(h_i, R)}{1-\delta} - c \geq \sum_{t\geq 1}\sum_{j\in\mathcal{I}} \delta^{t-1} u_j(\omega_t, P) \geq \frac{\sum_{j\in\mathcal{I}} \underline{u_j}(P)}{1-\delta},$$

where $\underline{u_j}(P)$ represents each player $j's$ minimum payoff at any state ω. Then, it must be that

$$c < \frac{\sum_{j\in\mathcal{I}} u_j(h_i, R)}{1-\delta} - \frac{\sum_{j\in\mathcal{I}} \underline{u_j}(P)}{1-\delta}.$$

A sufficiently high c violates this inequality.

Finally, consider two pairs (δ_1, c_1) and (δ_2, c_2) under which the optimal solution specifies never paying the switching cost c. Then, following (17), we can write:

$$c_k \geq \sum_{t\geq 1}\sum_{i\in\mathcal{I}} \delta_k^{t-1}\{u_i(\omega, R) - u_i(\omega, P)\}\pi(\omega, t), \quad k = 1, 2. \tag{18}$$

Multiply the first inequality for c_1 by λ and the second inequality for c_2 by $(1-\lambda)$ and we get that

$$\lambda c_1 + (1-\lambda)c_2 \geq \sum_{t\geq 1}\sum_{i\in\mathcal{I}} (\lambda\delta_1 + (1-\lambda)\delta_2)^{t-1}\{u_i(\omega, R) - u_i(\omega, P)\}\pi(\omega, t), \quad k = 1, 2, \tag{19}$$

which concludes our proof. ■

Proof of Proposition 7. The proof is by construction. We use the same parameters as in our benchmark simulations while adding the bipolar state. The transition in and out of this state is symmetric and equal to 0.45, whereas the utility of both countries choosing *power* is assumed to be 0.95. With these parameters, we obtain the following simulation:

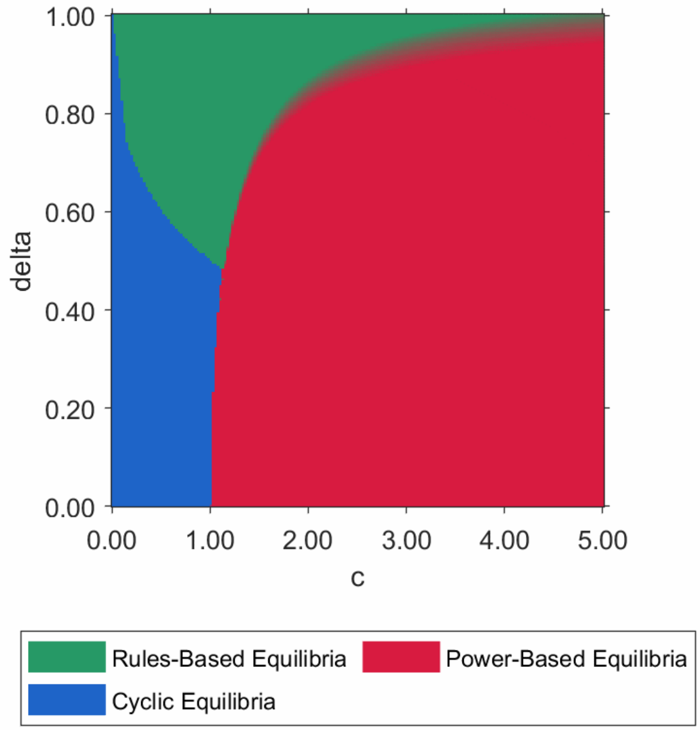

Figure 17: Rules-Based Equilibria in Bipolar World

This concludes the proof. ■

Proof of Proposition 8. Suppose that a rules-based equilibrium exists. Then, in the bipolar state it must be the case that

$$V_i(\Phi, R) \geq u_i(\Phi, (P, R)) + \delta \sum_{\omega'} V_i(\omega', P)$$

The condition above also holds for smaller values of $u_i(\Phi, (P, R))$. All other restrictions remain unchanged.

Relatedly, we also have that

$$u_i(\Phi, (P, R)) + \delta \sum_{\omega'} V_i(\omega', P) \geq u_i(\Phi, (P, R)) + \frac{\delta}{1-\delta}\underline{u}.$$

At the same time,

$$V_i(\Phi, R) \leq \frac{1}{1-\delta}\bar{u}.$$

Thus, a necessary condition for rules-based equilibrium is that:

$$u_i(\Phi, (P, R)) \leq \frac{\bar{u} - \delta \underline{u}}{1-\delta}.$$

Therefore, for sufficiently high $u_i(\Phi, (P, R))$, a rules-based equilibrium does not exist. ■

Proof of Proposition 9. Suppose, by contradiction, that there is a symmetric Markovian Equilibrium with $\sigma(\Phi_{ij}, \rho) = (R, R)$. Then, each country will have an incentive to unilaterally deviate and play P, getting a higher payoff at that period and not altering the status quo for the subsequent period. Moreover, due to symmetry, we cannot have $\sigma(\Phi_{ij}, \rho) = (R, P)$ or $\sigma(\Phi_{ij}, \rho) = (P, R)$, which concludes the proof. ■

Proof of Proposition 10. Suppose that there is an equilibrium in which both the hegemon and dominant countries choose to maintain R. Then, the following inequality must be satisfied:

$$u_i(d_i, R) + \delta \sum_{\omega'} q_{d_i,\omega'} V(\omega', R) \geq u_i(d_i, P) + \delta \sum_{\omega'} q_{d_i,\omega'} V(\omega', P) \tag{20}$$

Moreover, since the equilibrium choices are $\sigma(d_i, R) = R$ and $\sigma(d_i, P) = P$, we must have that:

$$V(d_i, R) = u_i(d_i, R) + \delta \sum_{\omega'} q_{d_i,\omega'} V(\omega', R), \tag{21}$$

$$V(d_i, P) = u_i(d_i, P) + \delta \sum_{\omega'} q_{d_i,\omega'} V(\omega', P), \tag{22}$$

which can be rewritten as:

$$V(d_i, R) = \frac{u_i(d_i, R) + \delta(q_{d_i,h_i} V(h_i, R) + q_{d_i,\Phi} V(\Phi, R))}{1 - \delta q_{d_i,d_i}},$$

$$V(d_i, P) = \frac{u_i(d_i, P) + \delta(q_{d_i,h_i} V(h_i, P) + q_{d_i,\Phi} V(\Phi, P))}{1 - \delta q_{d_i,d_i}},$$

Additionally, since in such equilibrium hegemons create R, we have that $V(h_i, P) = V(h_i, R) - c$ and given our result that under A2 there is no equilibrium in which both dominant countries play R in state Φ (Proposition 9), we have that $V(\Phi, R) = V(\Phi, P)$. Then, we can rewrite the condition above as:

$$V(d_i, R) = \frac{u_i(d_i, R) + \delta(q_{d_i,h_i}(V(h_i, P) + c) + q_{d_i,\Phi}V(\Phi, P))}{1 - \delta q_{d_i,d_i}}. \tag{23}$$

It must also be the case that $V(d_i, R) \geq V(d_i, P)$, since arriving at a state under regime R is always weakly better than arriving at it under regime P (Lemma 1, which also holds here), thus it must be true that

$$\frac{u_i(d_i, R) - u_i(d_i, P) + \delta q_{d_i,h_i} c}{1 - \delta q_{d_i,d_i}} \geq 0.$$

This can hold only if:

$$u_i(d_i, P) - u_i(d_i, R) \leq \delta q_{d_i,h_i} c. \tag{24}$$

Thus, a necessary condition for the existence of a pure strategy symmetric Markovian equilibrium in which hegemons create (and maintain) R and dominants maintain R is that:

$$q_{d_i,h_i} \geq \frac{u_i(d_i, P) - u_i(d_i, R)}{\delta c}.$$

This concludes the proof. ■

Proof of Lemma 2. Consider an equilibrium in which the dominant country chooses R when the status quo is R. If it does so under a strategy that specifies a rules breakdown, then

$$\begin{aligned} u_i(d_i, R) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', R) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, P)) \geq \\ u_i(d_i, P) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', P) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, P)). \end{aligned} \tag{25}$$

If it does so under a strategy that specifies a coordination on R, then

$$\begin{aligned} u_i(d_i, R) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', R) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, R)) \geq \\ u_i(d_i, P) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', P) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, R)). \end{aligned} \tag{26}$$

If it does so under a strategy that specifies a regime reversal, then

$$u_i(d_i, R) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', R) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, P)) \geq \tag{27}$$
$$u_i(d_i, P) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', P) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, R)).$$

If this is an equilibrium under a strategy that specifies the maintenance of the status quo, then

$$u_i(d_i, R) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', R) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, R)) \geq \tag{28}$$
$$u_i(d_i, P) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', P) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, P)).$$

Using a reasoning analogous to Lemma 1, it cannot be true that (25), (26) or (27) hold but (28) does not. ■

Proof of Proposition 11. Suppose that there exists a rules-based equilibrium in a bipolar world. This implies that

$$V_i(\Phi_{ij}, R) \geq u_i(\Phi_{ij}, (P, R)) + \delta \sum_{\omega'} V_i(\omega', P).$$

This condition implies that in the multipolar world under the strategy in which countries maintain the status quo at the multipolar state, countries also do not have an incentive to deviate from *rules* (recall that the static payoff in the situation where the country is the sole deviator to *power* under status quo *rules* is higher under bipolar than under the multipolar world). Thus, if there is a rules-based equilibrium in the bipolar world, there must be at least one rules-based equilibrium in the a multipolar world. ■

Proof of Proposition 12. Consider first a dominant country i that receives R as the status quo. In the short run, its incentive to choose P over R decreases, as $u_i(d_i, P) - u_i(d_i, R)$ falls. In the long run, in addition to remaining dominant, the country can become hegemonic or subordinate. As hegemonic, its gain with R relative to P increases. The same is true as subordinate. Therefore, the incentive of a dominant country to keep R rises. The analysis for a hegemonic country both for introducing and for keeping R is analogous. Therefore, if the payoff of all countries under R increases, a rules-based equilibrium becomes viable under a larger set of parameters. ■

Proof of Proposition 13. Let $\pi_{-i}(\omega, t)$ denote the probability that state $\omega \in \Omega \setminus \{h_i, d_i\}$ is reached at period t, without having reached states d_i or h_i in any period $t = 1, 2, ..., t-1$, where these periods are defined as t periods ahead of period 1 that starts at d_i. Conversely, denote by $\pi_i(\omega, t)$ the probability that $\omega \in \{h_i, d_i\}$ is reached in time t *and* neither d_i or h_i is ever reached in times $1, 2, ..., t-1$. Given that the transition probabilities are independent of actions and that this is a stationary world, we can write $V_i(d_i, RL)$ as:

$$\begin{aligned}
V_i(d_i, RL) &= u_i(d_i, RL) + \sum_{\omega \in \Omega \setminus \{h_i, d_i\}} \sum_{t>1} \delta^{t-1} u_i(\omega, RL) \pi_{-i}(\omega, t) + \\
&\quad \sum_{t>1} \delta^{t-1} V_i(d_i, RL) \pi_i(d_i, t) + \sum_{t>1} \delta^{t-1} V_i(h_i, RL) \pi_i(h_i, t) \\
V_i(d_i, RL)(1 - \sum_{t>1} \delta^{t-1} \pi_i(d_i, t)) &= u_i(d_i, RL) + \sum_{\omega \in \Omega \setminus \{h_i, d_i\}} \sum_{t>1} \delta^{t-1} u_i(\omega, RL) \pi_{-i}(\omega, t) + \\
&\quad \sum_{t>1} \delta^{t-1} V_i(h_i, RL) \pi_i(h_i, t) \\
V_i(d_i, RL) = \frac{1}{1 - \sum_{t>1} \delta^{t-1} \pi_i(d_i, t)} &\{u_i(d_i, RL) + \sum_{\omega \in \Omega \setminus \{h_i, d_i\}} \sum_{t>1} \delta^{t-1} u_i(\omega, RL) \pi_{-i}(\omega, t) + \\
&\quad \sum_{t>1} \delta^{t-1} V_i(h_i, RL) \pi_i(h_i, t)\}
\end{aligned}$$

Define $\kappa_\omega \equiv \sum_{t>1} \delta^{t-1} \pi_{-i}(\omega, t)$ for $\omega \in \{h_j, d_j\}$ and, similarly, $\kappa_\omega \equiv \sum_{t>1} \delta^{t-1} \pi_i(\omega, t)$ for $\omega \in \{h_i, d_i\}$. Then,

$$V_i(d_i, RL) = \frac{1}{1 - \kappa_{d_i}} \{u_i(d_i, RL) + \kappa_{h_i} V_i(h_i, RL) + \sum_{\omega \in \Omega \setminus \{h_i, d_i\}} \kappa_\omega u_i(\omega, RL)\} \qquad (29)$$

Moreover, due to our smoothness assumption, i.e., $q_{h_i,h_i} + q_{h_i,d_i} = 1$, $\forall i$, we have that:

$$\begin{aligned}
V_i(h_i, RL) &= u_i(h_i, RL) + \delta\{q_{h_i,h_i} V_i(h_i, RL) + q_{h_i,d_i} V_i(d_i, RL)\} \\
V_i(h_i, RL) &= \frac{1}{1 - \delta q_{h_i,h_i}} \{u_i(h_i, RL) + \delta q_{h_i,d_i} V_i(d_i, RL)\}
\end{aligned}$$

Thus, we can write equation (29) as:

$$V_i(d_i, RL) = \frac{1}{1-\kappa_{d_i}}\{u_i(d_i, RL) + \frac{\kappa_{h_i}}{1-\delta q_{h_i,h_i}}\{u_i(h_i, RL) + \delta q_{h_i,d_i} V_i(d_i, RL)\}+ \sum_{\omega\in\Omega\setminus\{h_i,d_i\}} \kappa_\omega u_i(\omega, RL)\} \tag{30}$$

Solving for $V_i(d_i, RL)$ leads us to:

$$V_i(d_i, RL) = \frac{1}{1-\frac{1}{1-\kappa_{d_i}}\frac{\delta\kappa_{h_i} q_{h_i,d_i}}{1-\delta q_{h_i,h_i}}} \frac{1}{1-\kappa_{d_i}}\{u_i(d_i, RL) + \frac{\kappa_{h_i}}{1-\delta q_{h_i,h_i}} u_i(h_i, RL)+ \sum_{\omega\in\Omega\setminus\{h_i,d_i\}} \kappa_\omega u_i(\omega, RL)\} \tag{31}$$

For ease of notation, define:

$$M_1 \equiv (1-\kappa_{d_i} - \frac{\delta\kappa_{h_i} q_{h_i,d_i}}{1-\delta q_{h_i,h_i}})^{-1}; \quad M_2 \equiv \frac{\kappa_{h_i}}{1-\delta q_{h_i,h_i}}; \quad M_3 \equiv \sum_{\omega\in\Omega\setminus\{h_i,d_i\}} \kappa_\omega.$$

Since $u_i(\omega, RL) = u_i(\omega, R) + \gamma$, if $\omega \in \{h_i, d_i\}$ and $u_i(\omega, RL) = u_i(\omega, R) - \lambda\gamma$, if $\omega \in \{h_j, d_j\}, j \neq i$, we have that:

$$V_i(d_i, RL) = M_1\{u_i(d_i, R) + \gamma + M_2(u_i(h_i, R) + \gamma) + \sum_{\omega\in\Omega\setminus\{h_i,d_i\}} \kappa_\omega u_i(\omega, R) - M_3\lambda\gamma\} \tag{32}$$

Maintaining RL in d_i is better than turning to P if $V_i(d_i, RL) \geq V_i(d_i, P)$, which is true if:

$$M_1\{u_i(d_i, R) + \gamma + M_2(u_i(h_i, R) + \gamma) + \sum_{\omega\in\Omega\setminus\{h_i,d_i\}} \kappa_\omega u_i(\omega, R) - M_3\lambda\gamma\} \geq V_i(d_i, P) \tag{33}$$

Note that for $\rho = P, R$, we can write:

$$V_i(d_i, \rho) = M_1\{u_i(d_i, \rho) + M_2 u_i(h_i, \rho) + \sum_{\omega\in\Omega\setminus\{h_i,d_i\}} \kappa_\omega u_i(\omega, \rho)\}$$

Therefore, from (33) we have that:

$$\gamma M_1(1 + M_2 - \lambda M_3) \geq V_i(d_i, P) - V_i(d_i, R) \tag{34}$$

By assumption of the statement in the proposition, we have that $V_i(d_i, P) - V_i(d_i, R) > 0$. Define the function $g(\gamma) \equiv \gamma M_1(1 + M_2 - \lambda M_3) - (V_i(d_i, P) - V_i(d_i, R))$. Note first that

($V_i(d_i, P)$ and $V_i(d_i, R)$) do not depend on γ, since this parameter is present only when the system is a rules-lite system, and does not affect the equilibrium payoffs of P or R. Moreover, $g(0) < 0$, since a rules-lite system with $\gamma = 0$ is the same as the rules-based system and since we assumed that $V_i(d_i, P) - V_i(d_i, R) > 0$. Conversely, if $\gamma = \bar{\gamma}$, then the rules-lite system yields a higher payoff for all players in all contingencies, implying that it must be better to sustain rules-lite rather than moving to power. Thus, $g(\bar{\gamma}) > 0$.

To conclude, we know that for any given set of parameters δ, Q, and $u's$, $g(0) < 0$ and $g(\bar{\gamma}) > 0$. Since $g(\gamma)$ is linear in γ with slope $M_1(1 + M_2 - \lambda M_3)$, there is a unique threshold $\gamma^\star$ as above. ■

References

D. Acemoglu and J. A. Robinson. Why did the west extend the franchise? Democracy, inequality, and growth in historical perspective. *The Quarterly Journal of Economics*, 115 (4):1167–1199, 2000.

P. Aghion, P. Antràs, and E. Helpman. Negotiating Free Trade. *Journal of International Economics*, 73:1–30, 2007.

K. Bagwell and R. Staiger. An Economic Theory of GATT. *American Economic Review*, 8 (1):215–248, 1999.

K. Bagwell and R. Staiger. Multilateral Trade Negotiations, Bilateral Opportunism and the Rules of GATT/WTO. *Journal of International Economics*, 63(1):1–29, 2004.

K. Bagwell and R. Staiger. Backward Stealing and Forward Manipulation in the WTO. *Journal of International Economics*, 82(1):49–62, 2010.

K. Bagwell and R. Staiger. Multilateral Trade Bargaining and Dominant Strategies. *International Economic Review*, 59(4):1785–1824, 2018.

K. Bagwell, R. Staiger, and A. Yurukoglu. Multilateral Trade Bargaining: A First Look at the GATT Bargaining Records. *American Economic Journal: Applied Economics*, 12(3): 72–105, 2020.

K. Bagwell, R. Staiger, and A. Yurukoglu. Quantitative Analysis of Multi-Party Tariff Negotiations. *Econometrica*, 89(4):1595–1631, 2021.

J. S. Becko, G. Grossman, and E. Helpman. Optimal tariffs with geopolitical alignment. *Mimeo*, 2025.

C. Boullenois, A. Kratz, and D. H. Rosen. Far from normal: An augmented assessment of China's state support. Technical report, Rhodium Group, 2025.

F. Broner, A. Martin, J. Meyer, and C. Trebesch. Hegemonic globalization. *CEPR Discussion Paper DP20339*, 2025.

S. G. Brooks and W. C. Wohlforth. *America Abroad. The United States' Global Role in the 21st Century*. Oxford University Press, 2016.

M. Camboni and M. Porcellacchia. Spheres of influence. *Mimeo*, 2025.

C. Clayton, M. Maggiori, and J. Schreger. A theory of economic coercion and fragmentation. *NBER Working Paper No. 33309*, 2024.

C. Clayton, M. Maggiori, and J. Schreger. A framework for geoeconomics. *Econometrica*, forthcoming.

I. Colantone, G. Ottaviano, and P. Stanig. The backlash of globalization. In *Handbook of international economics*, volume 5, pages 405–477. Elsevier, 2022.

E. Farhi and M. Maggiori. A Model of the International Monetary System. *Quarterly Journal of Economics*, 133(1):295–355, 2018.

R. C. Feenstra, R. Inklaar, and M. P. Timmer. The Next Generation of the Penn World Table. *American Economic Review*, 105(10):3150–3182, 2015.

C. Freund and E. Ornelas. Regional Trade Agreements. *Annual Review of Economics*, 2: 139–167, 2010.

P. Goldberg and T. Reed. Is the global economy deglobalizing? *Brookings Papers on Economic Activity Spring*, (2), 2023.

J. Goldstein and J. Gowa. US national power and the post-war trading regime. *World Trade Review*, 1(2):153–170, 2002.

G. Grossman and A. Sykes. Industrial policy and subsidies: Assessment of current rules and possible reforms. *Journal of World Trade*, 59(4):531–578, 2025.

H. Horn and P. C. Mavroidis. Why the US and the WTO should part ways. *VoxEU (25 June*, 2025.

D. A. Irwin. *Free Trade Under Fire.* Princeton University Press, 2020.

R. O. Keohane. The Demand for International Regimes. *International Organization*, 36(2): 325–355, 1982.

R. O. Keohane. *After Hegemony: Cooperation and Discord in the World Political Economy*. Princeton University Press, 1984.

C. Kindleberger. *The World in Depression, 1929-1939*. Berkeley and Los Angeles: University of California Press, 1973.

C. Kindleberger. The Rise of Free Trade in Western Europe, 1820-1875. *The Journal of Economic History*, 35(1):20–55, 1975.

M. J. Klasing and P. Milionis. Quantifying the Evolution of World Trade, 1870–1949. *Journal of International Economics*, 92(2):185–197, 2014.

B. Kleinman, E. Liu, and S. J. Redding. International friends and enemies. *American Economic Journal: Macroeconomics*, 16(4):350–385, 2024.

S. D. Krasner. State power and the structure of international trade. *World Politics*, 28(3): 317–343, 1976.

A. O. Krueger. Trump's tariff chaos could reverse 80 years of economic progress. *Project Syndicate (23 April)*, 2025.

N. Limao. Preferential Trade Agreements. In K. Bagwell and R. W. Staiger, editors, *Handbook of Commercial Policy*, volume 1-B, pages 279–367. North-Holland, Amsterdam, 2016.

R. D. Ludema and A. M. Mayda. Do Terms-of-Trade Effects Matter for Trade Agreements? Theory and Evidence from WTO Countries. *Quarterly Journal of Economics*, 128(4): 1837–1894, 2013.

G. Maggi. The Role of Multilateral Institutions in International Trade Cooperation. *American Economic Review*, 89(1):190–214, 1999.

G. Maggi. International trade agreements. In *Handbook of international economics*, volume 4, pages 317–390. Elsevier, 2014.

G. Maggi and A. Rodriguez-Clare. The value of trade agreements in the presence of political pressures. *Journal of Political Economy*, 106(3):574–601, 1998.

G. Maggi and A. Rodríguez-Clare. A political-economy theory of trade agreements. *American Economic Review*, 97(4):1374–1406, September 2007.

A. Mattoo and R. W. Staiger. Trade Wars: What do they mean? Why are they happening now? What are the costs? *Economic Policy*, 35(103):561–584, 2020.

A. Mattoo, M. Ruta, and R. Staiger. Geopolitics and the World Trading System. *NBER Working Paper No. 33293*, 2025.

P. C. Mavroidis. *The WTO Dispute Settlement System. How, Why and Where?* Northampton: Edward Elgar Publishing, 2022.

P. C. Mavroidis. *End of the Line? Industrial Policy, National Security, and the Perilous Plight of the WTO.* New York: Oxford University Press, 2025.

J. McLaren. Size, Sunk Costs, and Judge Bowker's Objection to Free Trade. *American Economic Review*, 87(3):400–420, 1997.

A. Nicita, M. Olarreaga, and P. Silva. Cooperation in WTO's Tariff Waters? *Journal of Political Economy*, 126(3):1302–1338, 2018.

C. Norrlof. *Hegemony*. New York: Oxford University Press, 2015.

P. K. O'Brien and G. A. Pigman. Free Trade, British Hegemony and the International Economic Order in the Nineteenth Century. *Review of International Studies*, 18(2):89–113, 1992.

R. Ossa. Trade Wars and Trade Talks with Data. *American Economic Review*, 104(12): 4104–4146, 2014.

R. Ossa. Pillars of Economic Security. *Finance & Development*, IMF, 2023.

A. Peres. World Trade Organization: Challenges and opportunities. *Research Briefing, House of Commons Library*, 2024.

G. A. Pigman. Hegemony and Free Trade Policy: Britain 1846-1914 and U.S.A. 1944-1990. *PhD thesis, University of Oxford*, 1992.

G. A. Pigman. Hegemony and Trade Liberalization Policy: Britain and the Brussels Sugar Convention of 1902. *Review of International Studies*, 23(2):185–210, 1997.

D. Rodrik. How to smooth US-China economic relations for the benefit of the global economy: A light model of global economic governance. *Journal of Government and Economics*, 13: 100097, 2024.

L. S. Shapley. Stochastic Games. *Proceedings of the National Academy of Sciences of the United States of America*, 39(10):1095–1100, 1953.

D. Snidal. The Limits of Hegemonic Stability Theory. *International Organization*, 39(4): 579–614, 1985.

R. Staiger. *A World Trading System for the Twenty-First Century*. Dartmouth and NBER, 2022.

J. Stiglitz and D. Rodrik. Rethinking global governance: Cooperation in a world of power. *Mimeo*, 2024.

M. C. Webb and S. D. Krasner. Hegemonic Stability Theory: An Empirical Assessment. *Review of International Studies*, 15(2):183–198, 1989.

B. Zissimos. The End of the End of History: A Political-Economy Perspective. *Intereconomics*, 57(6):372–376, 2022.