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

Cecilia Carvalho† Daniel Monte‡ Emanuel Ornelas§

March 20, 2025

**Abstract**

The rules of the World Trade Organization are increasingly being disregarded by its members, raising concerns about the future of multilateral liberalization. In this paper, we develop a dynamic model of the international trading system to analyze the sustainability of rules-based trade regimes. We consider a framework of stochastic asynchronous games, where a leading country determines the trade regime and the identity of this leader changes over time. In a many-country, infinite-horizon game, we show that transitioning from a power-based to a rules-based regime requires the presence of a hegemonic power – i.e., a country significantly larger than all others. Dominant countries benefit from a power-based regime but may choose to maintain rules because of the possibility of losing their dominance in the future. We find that the long-term viability of a rules-based equilibrium hinges on the cost of establishing it, which must be neither too small nor too large. If the cost is too small, the system follows a cyclic equilibrium; if too large, a power-based regime prevails. In a bipolar state, free-riding and market-power forces further undermine the feasibility of rules-based equilibria. Our findings highlight the risks posed by the ongoing erosion of multilateral cooperation and its potential long-term consequences for the world trading system.

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

**JEL classification:** F02; F13; F53

*We are grateful to Kyle Bagwell, Stephen Brooks, Nuno Limao, Tom Prusa and Gregory Shaffer for detailed discussions of our paper. We also thank Paula Bustos, Camila Campos, Andrés Carvajal, Paola Conconi, James Fenske, Petros Mavroidis, Isabelle Mejean, Joao Paulo Pessoa, Dani Rodrik, David Rubin, Bob Staiger, Ben Zissimos, and seminar participants at the 2024 GEP/CEPR Postgraduate Conference, 2024 Gravity, Trade Agreements and Policy Workshop, 2024 Rethinking the Global Trading System Workshop, *Staigerfest* Workshop, 2024 TIGN Workshop, 2023 InsTED Workshop, 2023 World Congress of the International Economic Organization, EPGE, Geneva Trade and Development Workshop, INSPER, Online Australasian Seminar in International Economics, Paris School of Economics, Sao Paulo School of Economics and USP for valuable comments and discussions. We thank Nicolas Moura for excellent research assistance. Ornelas acknowledges financial support from the Brazilian Council of Science and Technology (CNPq).

†University of Geneva (GSEM) and HES-SO/HEG-Genève. E-mail:ceciliambc@gmail.com.

‡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 order has operated under a rules-based system 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, waged mainly (but not only) between the U.S. and China, 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. no longer appears willing to sponsor 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 this possible change and its 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 country, and we assume that the identity of the largest country changes stochastically. We show that when the cost to establishing a rules-based system is small, the world experiences a cyclic equilibrium, transitioning between rules-based and power-based regimes. If that cost is large, then the world gets stuck in a power-based equilibrium, even if collectively countries would benefit from a rules-based regime. It is only when the cost to establishing a rules-based system is moderate and discount factors are sufficiently high that such a system can be sustained in equilibrium. 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 study the long-run viability of trading regimes, we follow the established literature to define countries' static payoff ordering. Specifically, we assume that in the short run the "leader" country would prefer a power-based regime *if* all other countries participated. This would allow the leader to obtain the highest gains from trade, taking advantage of its size in the negotiations. Thus, "dominant" countries prefer to operate under a power-based system. However, if size asymmetries are very large, i.e., if the leading country were not only dominant but a "hegemon," then other countries would not want to participate in a power-based regime.[1] As a result, the hegemon, unlike the dominant, is better off under a rules-based system, because it ensures participation of smaller countries.

[1]This can be rationalized, for example, when production requires irreversible investments and the private sector concentrates investment in export-oriented sectors in anticipation of a liberal trade regime (as, e.g., in McLaren, 1997). This would reduce the (ex-post) bargaining power of small countries. Anticipating that, under a power-based system small countries decide not to engage in trade deals at all.

These scenarios have been recently discussed by Mattoo and Staiger (2020) and Staiger (2022). Those static regime preferences also 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[2] 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 – a country significantly larger than all other countries. 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. This yields two main conclusions. First, the presence of a hegemon fosters the establishment and maintenance of a rules-based international liberal regime. Second, smaller nations, not bearing the costs but sharing the benefits from this arrangement, gain as well under hegemonic rule. This enhanced stability yields collectively desirable results for all countries in the international system. However, it relies on the existence of a hegemon; in the absence of hegemonic power, global disorder and inefficient outcomes prevail. The theory therefore predicts that a rules-based trade regime will endure only while there is a hegemon to sponsor it.[3] Based on these insights, the HST associates the establishment and sustainability of liberal trading regimes in the mid-19th and since the mid-20th centuries with the British and American hegemonic leaderships, respectively, and the shift to protectionism between world wars with the lack of a hegemon.

This is only a static perspective, however. It 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 easy, but (re-)establishing a rules-based regime is costly for the leader, who must bear the coordination and implementation burdens. The leader's decision 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 envi-

[2]Keohane (1982) refers to this public good as a "regime," encompassing several international issue areas such as trade, monetary relations, energy, foreign investment, and security.

[3]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. By extending their reach, our model formalizes some of HST's results, qualifies others, and pushes their scope significantly further.

ronment 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 better captures 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 a non-hegemonic dominant country that inherits a power-based regime will always keep it as well. Again, we show that there is no long-run gain from replacing it for a dominant country, even though that country would be better off in the future, when it is no longer a leader, under a rules-based regime. 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, in any symmetric Markovian equilibrium, a rules-based trading system can emerge *only* when there is a hegemon to sponsor it.

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 do bear it, and the world will be permanently under a power-based regime. Thus, the existence and endurance of a rules-based system requires the cost to reestablish it to be neither too large nor too small. Furthermore, countries must be sufficiently forward-looking, or else the equilibrium will be either cyclic or power-based.

We show that the decentralized equilibrium is often inefficient in that it fails to maximize the present value of the world's aggregate payoff. Notably, this inefficiency systematically biases the system against sustaining a rules-based equilibrium. 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 type of inefficient power-based equilbria arises due to a collective action problem, in the sense that, if a rules-based regime were put in place, it would persist. The problem is that no leading country is 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 then relax this assumption by allowing dominant leaders to prolong their dominance when choosing

a power-based regime. While this introduces an additional consideration, it does not alter the fundamental tradeoffs of the benchmark model. Interestingly, this change may not even make a rules-based equilibrium less likely. Although the direct "endogenous dominance effect" makes a power-based regime more appealing for dominant countries, there are indirect effects that can offset it. In fact, we show that a rules-based system remains possible even in the extreme case where a dominant country can secure indefinite leadership by choosing *power*. The reason is that dominant countries recognize the possibility of becoming a hegemon in the future, and in that state they are better off under *rules*. Anticipating this, they may prefer to preserve the rules-based system rather than incurring the cost of reintroducing it later, even if doing so increases the risk of eventually losing leadership.

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 of course 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 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 "plays by the rules," i.e., chooses trade barriers without attempting to exploit its own market power. In this context, we identify three 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 in 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 could completely undermine the sustainability of a rules-based equilibrium. We conclude that sustaining rules-based equilibrium in a bipolar state poses critical challenges, stemming from free-riding and market-power motives.

Second, we introduce the possibility of a leaderless state, where all countries are "subordinate." In this scenario, establishing *rules* requires unanimity and imposes a cost on each country. While unanimity is not strictly necessary, it captures the idea that, in the absence of a dominant leader, a cooperative regime requires effort from multiple players. Consistent with the literature, we assume that each country in a leaderless state prefers *rules* over *power*. We show that it is easier to sustain a rules-based equilibrium when the world has such a leaderless state than when the world has a bipolar state. The main reason is that, in the former, market-power incentives are absent.

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 what would be the consequences for the multilateral trading system. Mattoo and Staiger (2020) 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, with a significant risk of 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 what we do in this paper. Specifically, we rely on their insights to motivate some key building blocs 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, this type of analysis is entirely novel in 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.[6]

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.[7] More recently, quantitative models have provided a better understanding of the magnitudes involved. Ossa (2014) shows that noncooperative tariffs, based on power, lead to a large drop in trade volumes, along with a decrease in overall welfare. As a result, nobody wins in the trade war. In turn, the inefficiency of these tariffs creates incentives for multilateral trade negotiations. He

[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. (2024) 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 this paper, they study how hegemons affect economic integration. However, our model differs from theirs in many dimensions: the questions addressed, the modeling approach, and the insights obtained.

[6]This is reflected in the rising number of both policy articles (see, e.g., Goldberg and Reed, 2023), and of newspaper articles (see, for example, the May 11th 2024 issue of The Economist, whose cover was "The New Economic Order") on "deglobalization" and related phenomena in recent years. Those concerns have been vastly amplified after the 2024 election of Donald Trump in the U.S..

[7]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, by internalizing these effects, they result in lower average applied tariffs compared to noncooperative levels. 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. (2024) show, using a sample spanning 180 years, that hegemons are associated with more – and more effective – trade agreements.

finds that cooperative tariffs, gradually negotiated, yield large welfare gains. Naturally, a rules-based system has value only if there is a mechanism to enforce the rules. The 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.

We see our contribution as an initial step toward understanding the future of the rules-based world trading system. Inevitably, the analysis is stylized in many ways. For example, we abstract from geopolitical rivalry related to countries' relative economic positions (Mattoo et al., 2024) and from concerns over economic security (Clayton et al., 2024), which may be affected by the trade regime. 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 them 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. Another simplifying assumption of our analysis is that we take for granted the existing set of "rules" of the international trading system. However, in the conclusion we indicate, in light of our model, the key tradeoffs that policymakers should consider when redesigning the current system of rules toward making it more sustainable.

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, as well as its current challenges. In section 3, we develop the benchmark model. In section 4, we endogeneize 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 conclude in section 7.

# 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 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 and 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. 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 rule 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 agreement 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]

[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.

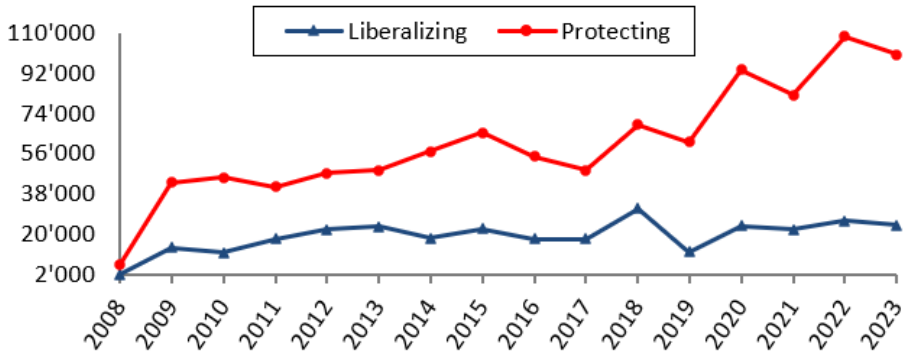



Figure 1: World trade policy interventions (2008 - 2023)

Source: Global Trade Alert database, 2024

## 2.2 The Erosion of the Rules-Based System

Despite its historical success, the GATT/WTO trading system has faced a set of challenges in recent decades, which 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.

We have observed increasing disregard for the GATT/WTO system in recent years. The following two figures illustrate this trend 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]

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 two largest trading

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 liberalises on a non-discriminatory basis or improves the transparency of a relevant policy. In our data, we merge the first and the second evaluations into the category of "protecting," while classifying the third as "liberalizing." In the database, a single state act can result in one 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 it would follow the same pattern.

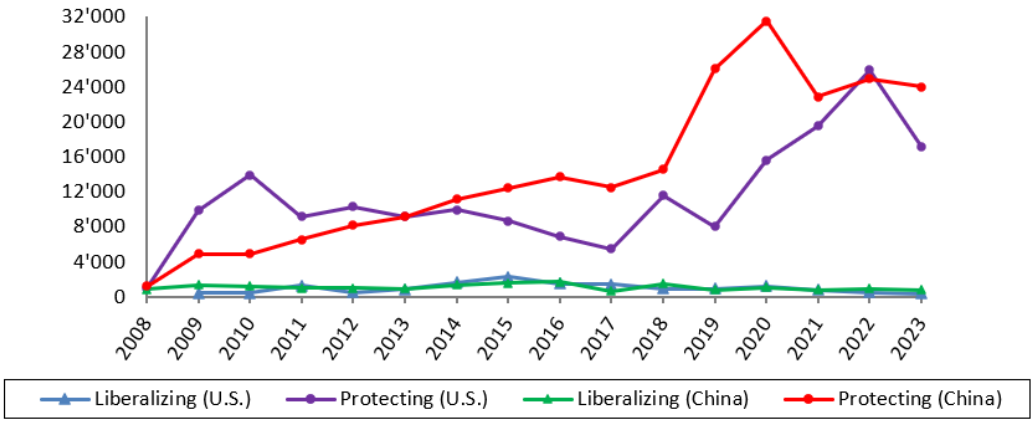



Figure 2: U.S. and China trade policy interventions (2008 - 2023)

Source: Global Trade Alert database, 2024

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.

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

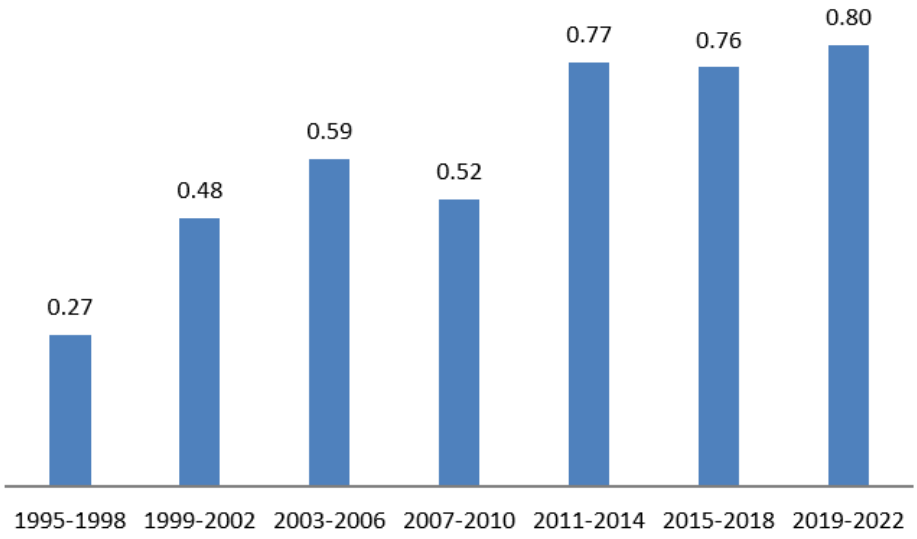



Figure 3: Average of DSP consultations moved to panels (1995 - 2022)

Source: WTO Annual Report, 2023

difficulty for WTO members in resolving disputes solely through the consultation mechanism.[11]

Now, if a panel is formed and 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 fewer than three members (it currently has none), thereby becoming unable to consider any case. 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.

Staiger (2022) highlights other challenges for the rules-based multilateral trading system (e.g., the rise of large emerging markets, digital trade, and the rise 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.

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 power over that period. In parallel, in the last decades the world witnessed the rise of China, which became the second largest economy in the

[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 of the implications of this behavior.

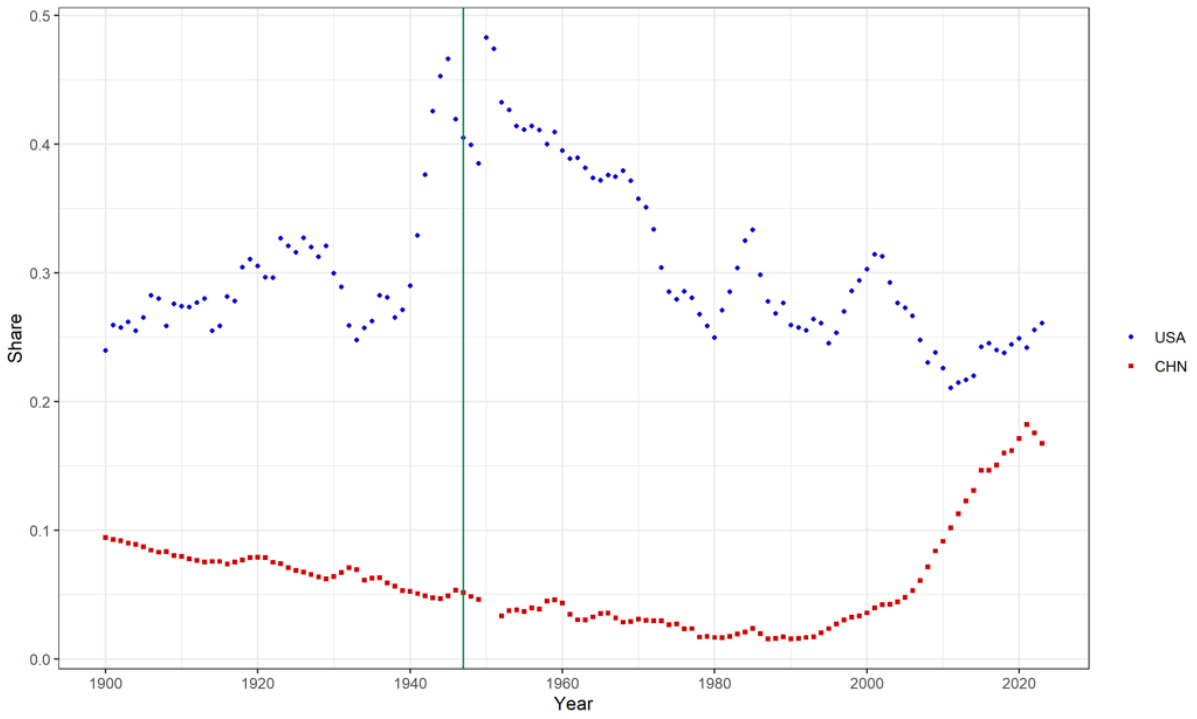



Figure 4:
Share of world GDP (1900 - 2023, Current Prices)

Source: 1870-1949: data from Klasing and Milionis, 2014; 1950-1959 data from Penn World Tables 10.01 (Feenstra et al., 2015); 1960-2023 data from World Bank, 2025

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, with its share of world GDP increasing significantly, particularly 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 entered into 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 a leading economic force, the rise of China means that it no longer holds the unchallenged dominance it once did.

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 moved 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 reciprocity and nondiscrimination as described in the GATT/WTO, and rests instead on countries' bargaining power.

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 $\omega$ a typical element of $\Omega$.

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.[12] Specifically, in states $h_i$ and $d_i$ country $i$ chooses a trade regime $\rho$ 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 at those states. At any state, one country is the leader (either hegemon or dominant) and $N - 1$ countries are subordinate.

To define the ordering of countries' payoffs under each regime, we follow Mattoo and Staiger (2020). They 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. Specifically, 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 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, its static gains are higher under regime $R$ than under regime $P$. Moreover, when a country is subordinate, it always prefers regime

[12] 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 of 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.

$R$. In turn, the second inequality says that when $i$ is dominant, its static gains are higher under regime $P$.

The rationalization of this ordering is as follows. 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 the expense of the smaller countries. 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 of the ground.[13]

We define a country as a *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 – as 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 define a country as *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.[14] In Appendix C, we elaborate further on the ordering of payoffs under each state.

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

[13]In a similar vein, Goldstein and Joanne (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. 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 futures expropriation. Unlike those papers, here the commitment (offered by a rules-based regime) is useful in a static sense, but in the long run.

[14]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).

- *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.

## 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 be $(\omega_0, \rho_0)$. Since we are concerned with a model that can account for the creation of the rules-based regime, we consider the case in which $\rho_0 = P$.[15] 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 $\omega$ to state $\omega'$. 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. In the benchmark model, we take matrix $Q$ as exogenous. In section 4, we allow the choice of regime to affect the transition probabilities.

More precisely, here we consider a situation where if a country is 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.

[15] 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 is always chosen on the equilibrium path even if no country is willing to incur the cost to introduce it.

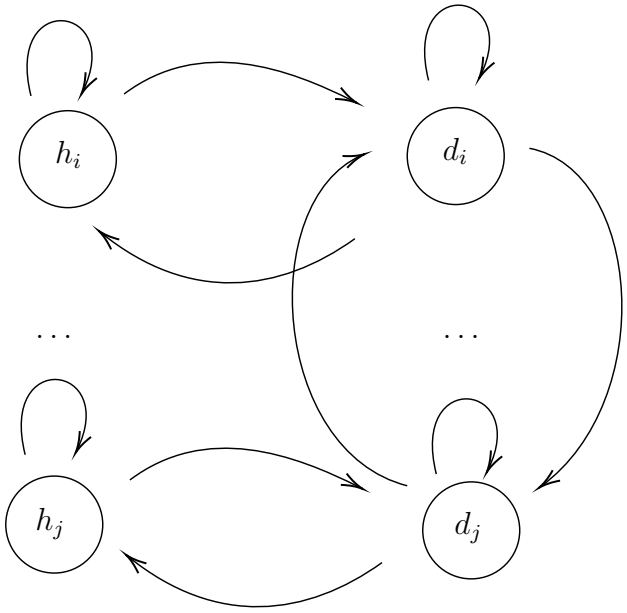



Figure 5: Transition between states

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. However, 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 enforce it, or any multilateral coordination.[16]

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*

[16] 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.

*choice in state $\omega$ when the status quo regime is $\rho$. 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.*

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 $\rho \in \{P, R\}$. At states $h_i$ and $d_i$, country $i$'s payoff functions depend on the inherited and chosen regimes. For $\omega \in \{h_i, d_i\}$ :

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

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:

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

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 3.4. 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 D.

**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, there is always a cyclic equilibrium (and it is the unique Markovian symmetric 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 Markovian symmetric 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 system 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 as dominant leader, the regime is $P$; when the world has a hegemon leader, 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 be sustained, as the following proposition shows.

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

A. *There exists $\underline{c} > 0$ such that, if $c < \underline{c}$, a rules-based equilibrium does not exist.*

B. *There exists $\bar{c} > \underline{c}$ such that, if $c > \bar{c}$, a rules-based equilibrium does not exist.*

C. *There exists $\underline{\delta} > 0$ such that, if $\delta < \underline{\delta}$, a rules-based equilibrium does not exist.*

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 never being their most advantageous static choice.

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 a 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 directly 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 $R$. This is because the cost $c$ may discourage a country from switching from regime from $P$ 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 faced with 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, note that in any pure strategy symmetric Markovian equilibrium either $\sigma(d_j, P) = P$ or $\sigma(d_j, P) = R$, $\forall j \in \{1, 2, ..., N\}$. Suppose, by contradiction, that it is the second case, i.e., $\sigma(d_j, P) = R$. First, note that when a dominant country inherits a power-based regime, in the short run it is better-off maintaining $P$, by assumption. 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 may (i) remain dominant, (ii) become hegemon, or (iii) another country may become dominant. If the country remains dominant in the subsequent period, the country would have benefited from keeping $P$. If the country becomes a hegemon, it would again have benefited from keeping $P$ (and potentially choosing $R$ only after becoming a hegemon). Finally, if the country 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 $\sigma(d_j, P) = P \ \forall j \in \{1, 2, ..., N\}$.

Proposition 3 has important implications. The first is useful from a technical perspective, as 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). Using the results of the proposition, we reduce the set of possible pure strategy symmetric Markovian equilibria to only four possibilities: rules-based, power-based (which can be of two subclasses) or cyclic equilibria (see Definition 3).[17]

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

The 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 any 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.*

In fact, hegemons are not only indispensable: if they are sufficiently rare, 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.*

Hence, establishing a rules-based system requires a hegemon and the world needs "enough" hegemons (in a probabilistic sense) to ensure it is maintained in the long run. In this sense, Corollary 2 and Proposition 4 offer support for the main premise of Kindleberger's (1973) Hegemonic

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

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, they do not imply that a hegemon is needed to *keep* a rules-based system in place.

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 $R$, the rules-based equilibrium exists only in mixed strategies. A rules-based equilibrium in mixed strategies is one in which play at all states $h_i$, $\forall i$, involve 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 an 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 $P$, $\bar{c}$, there is a discontinuity. This happens because we are focusing on symmetric equilbria, so that, for $c$ just below $\bar{c}$, 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 $P$.

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 subsection 3.1. For simplicity, we consider a symmetric transition matrix $Q$.[18] 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*.[19] We then check for which values of $(\delta, c)$ the system has a rules-based symmetric Markovian equilibrium. The 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.

[18] 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 $i \neq j \neq k \neq l$.

[19] 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 $q_{h_1,d_2} = q_{h_1,h_2} = q_{d_1,h_2} = q_{d_2,h_1} = q_{h_2,h_1} = q_{h_2,d_1} = 0$. 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 county obtains a larger share of the gains than the subordinate country.

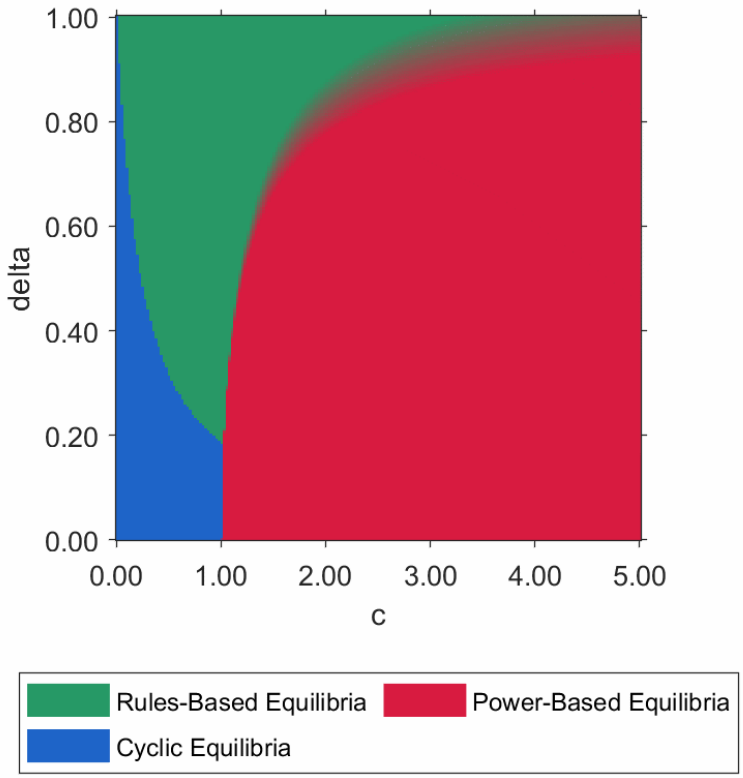



Figure 6: Symmetric Markovian equilibria

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 $\delta$ 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 $\delta$, 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 $\delta$ that defines the region where there is a rules-based equilibrium becomes positive after that point – that is, with a higher $\delta$, 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 $\delta$), then a rules-based equilibrium cannot exist (Proposition 1.C). Finally, the area corresponding to the shadowed region between the rules-based (green) and power-based (red) regions reflects equilibria with mixed strategies. 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 $P$.

To illustrate Proposition 4, we consider the same baseline parameters of 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 $\delta$ that allows for a rules-based equilibrium rises, in line with Proposition 4. If that probability reaches 0.99, the possibility of such an equilibrium virtually vanishes, even for $\delta$ very close to one.

It is worth noting 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.[20] 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:

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

It can be shown that all previous results hold in this world with $\varepsilon$-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*.

[20] Indeed, Mavroidis (2025) argues that the current WTO rules are outdated. This is also a central message of Mattoo et al. (2024), but due to geopolitical rivalry motives.

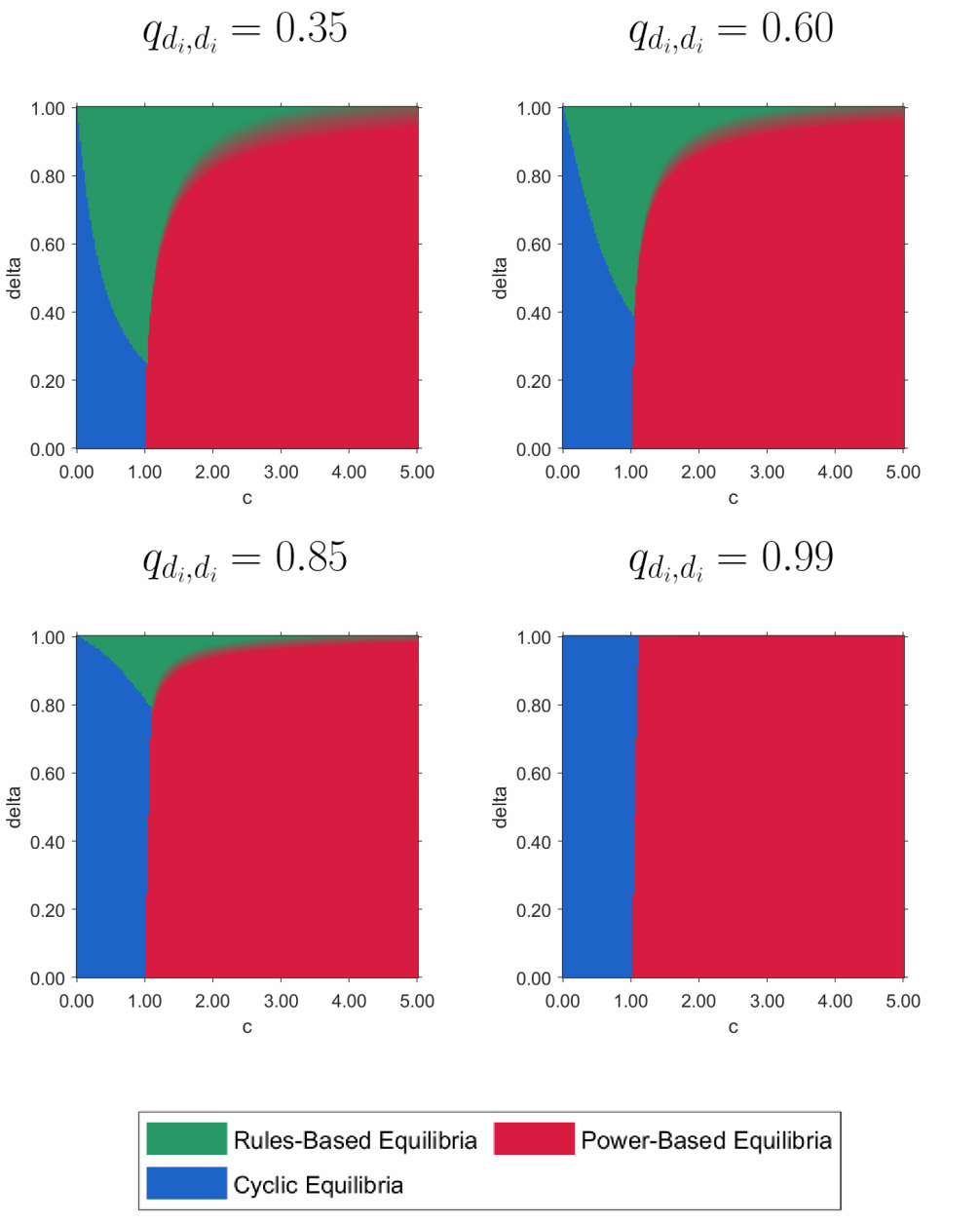



Figure 7: Enduring dominance

### 3.3 Efficiency

An important question is whether/when inefficiencies arise in the choice of international trade regimes. For example, 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. We keep the assumption that the initial state $(\omega_0, \rho_0)$ is exogenously set at $(\omega_0, \rho_0) = (h_1, P)$. 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 as we have done in our 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 the most 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 $(\delta, c)$ with $\delta > \delta(c)$, the efficient solution is rules-based. In contrast, for any pair $(\delta, c)$ with $\delta < \delta(c)$, the efficient solution is power-based.

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

**Proposition 5 (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 $R$. That is, 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 $(\delta', c')$ with $\delta' \leq \delta$ and $c' \geq c$ also belongs to $S$.*
2. *For any $(\delta, c), (\delta', c') \in S$, then $\lambda(\delta, c) + (1-\lambda)(\delta', c')$, $\forall\lambda \in (0, 1)$, also belong to $S$.*

Proposition 5 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 would

choose *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 above, 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 (the cyclic equilibria), knowing that hegemons will revert back to rules. 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.

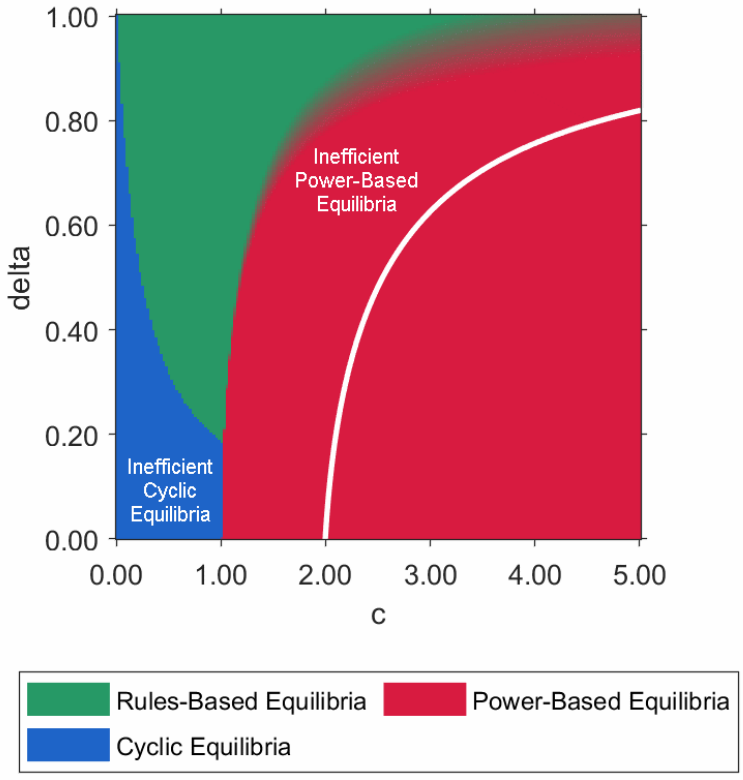



Figure 8: Efficient and inefficient equilibria

## 3.4 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 cost. 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 4 (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. Moreover, if a dominant country receives $R$ as the status quo (which must be off-equilibrium, since $\rho_0 = P$), it will switch to $P$.

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

**Definition 5 (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 $R$. This can be seen more clearly in Figure 9, which 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 $c$.

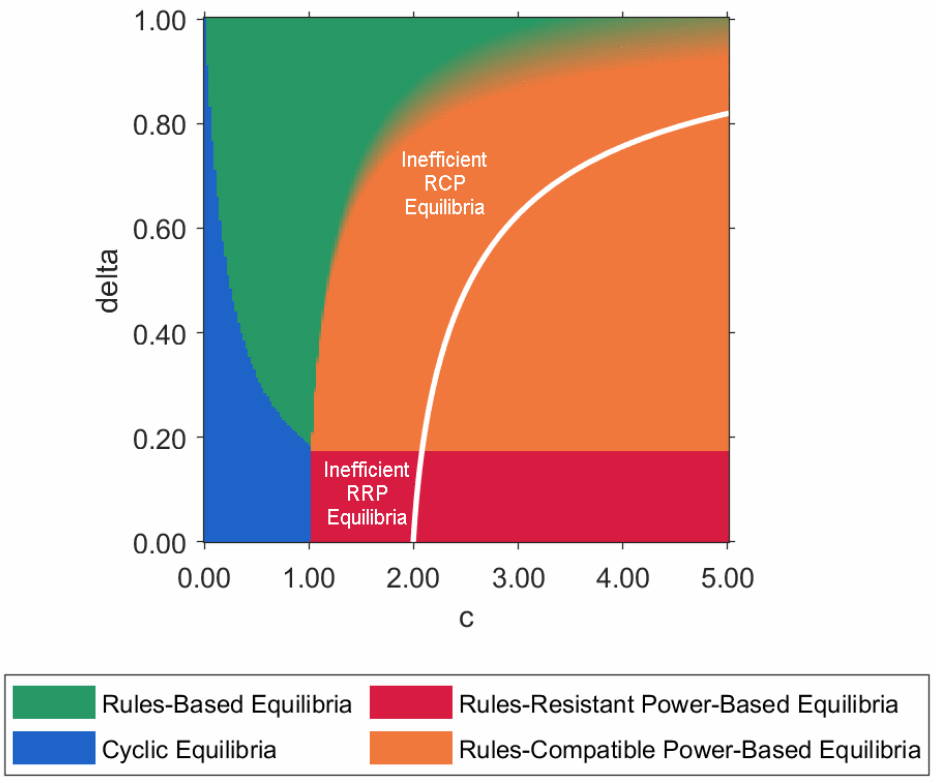



Figure 9: Rules-resistant and rules-compatible power-based equilibria

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.

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.

# 4 Endogenous Transitions

We have thus far 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. 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 remaining a dominant country is more likely if the dominant country chooses $P$. That is, for any given initial trade regime $\rho$:

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

for $\gamma \geq 0$.[21] 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 a different one. Our analysis so far has assumed $\gamma = 0$; we are now interested in the case where $\gamma > 0$. This is the most natural case to consider for endogenous transition probabilities. Under a power-based regime, larger countries obtain (static) payoffs more aligned with their economic sizes. Presumably, this can make the power asymmetries last for a longer period, relative to the benchmark case in which transition probabilities are exogenous.

It is relatively 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 $\gamma$.

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

However, there are two opposing indirect effects at work as well. First, as $\gamma$ rises, the probability that a dominant leader remains in power 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 selecting $P$, the same logic applies to country $j \neq i$. This implies that losing leadership when the prevailing regime is $P$ becomes increasingly costly, as the new dominant will choose $P$ if the status quo is $P$. Thus, while *power* helps sustain dominance, it also makes the prospect of losing leadership more detrimental, which discourages the dominant country from selecting *power*.

It then follows that, even if perhaps counter intuitively at first, allowing for $\gamma > 0$ does *not* necessarily make a rules-based equilibrium less likely. The net effect of increasing $\gamma$ is, in general, ambiguous. For example, using the same baseline parametrization as Figure 6, we find that increasing $\gamma$ has minimal impact on the equilibrium regions. Figure 10 illustrates this point in the starkest way, by considering the $\bar{\gamma}$ that entirely shuts down the possibility of leadership loss when a dominant country chooses $P$ – meaning that, if a dominant country chooses $P$, it remains leader, whether dominant or hegemon, with certainty. In this scenario, the second indirect effect disappears, eliminating strategic interactions. Yet, the conditions for a rules-based equilibrium remain largely unchanged compared to Figure 6. In fact, when $c$ and $\delta$ 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 $R$.

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.

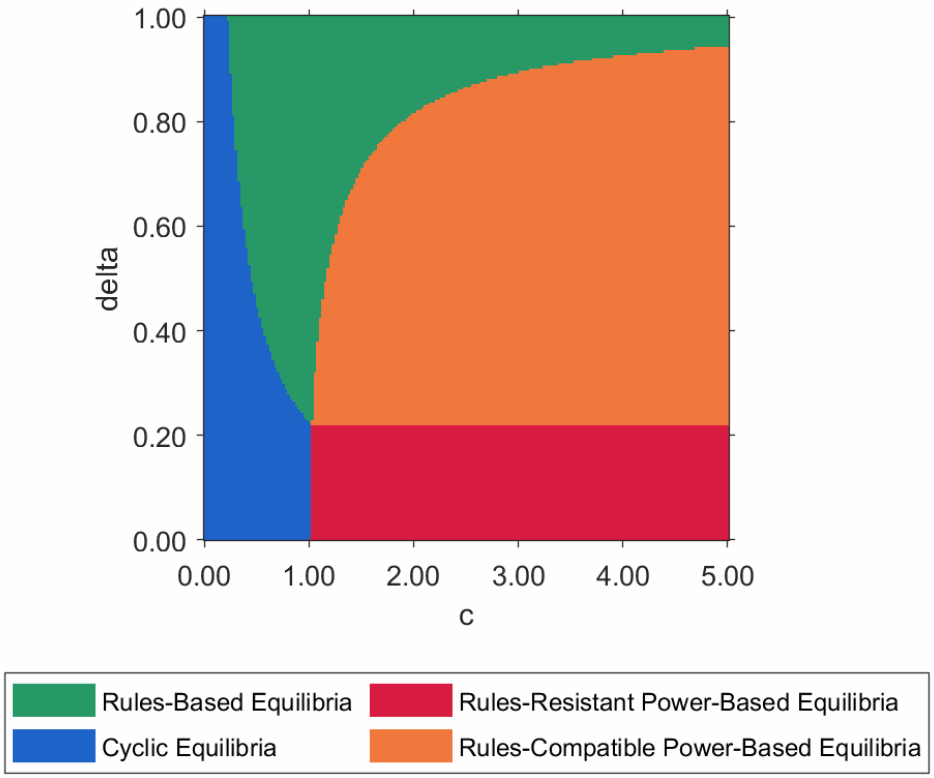



Figure 10: Equilibria under Endogenous Transitions, $\gamma = \bar{\gamma}$

# 5 A Bipolar World

We now extend the model to incorporate a state in which there are two leader 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 $\Phi_{ij}$. Thus, in a two-country version, we now have that $\Omega = \{h_1, d_1, \Phi_{12}, d_2, h_2\}$. We maintain the assumption that the transitions are smooth, that is, all transitions to neighbor states are positive and $\forall i$, $q_{h_i,h_i} + q_{h_i,d_i} = 1$, $q_{d_i,h_i} + q_{d_i,d_i} + q_{d_i,\Phi} = 1$ and $q_{\Phi,d_1} + q_{\Phi,\Phi} + q_{\Phi,d_2} = 1$. Figure 11 illustrates this situation.

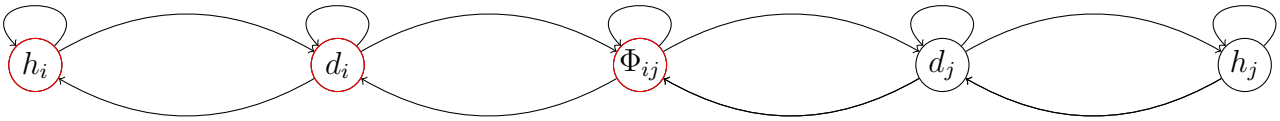



Figure 11: Transition between states; bipolar world

Whenever the system reaches state $\Phi_{ij}$, the two dominant countries 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 $(\rho_i, \rho_j)$:

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

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, but if one plays *rules* while the other dominant country does not, exploiting its market power, then the former would be in a worse situation than it would have been had it abandoned *rules* itself.[22]

[22]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, 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 6 (Rules-Based Equilibrium if 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 multiple 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 7 (Market Power Undermines Rules-Based Equilibria)** *Assume A1 and suppose that there is a rules-based equilibrium. In particular, at state $\Phi_{ij}$ 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 7 shows that large short-run gains from exploiting market power in international markets in a bipolar state can be critical to undermining a rules-based equilibrium's viability.

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 (namely, the need for joint support for *rules* instead of a single leader's support). However, precisely because under

A2 free-riding is not consequential 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 $\Phi_{ij}$, a world under a rules-based regime will switch to a power-based system.

**Proposition 8 (Bipolar State Begets Power with Positive Probability)** *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 8 still holds, highlighting the fragility of rules-based equilibria.

To conclude this section, we present a result that relates HST to the context of a world with a bipolar state. Namely, in a world in which there is a bipolar state and assumption A2 holds, dominant countries will sustain $R$ only if the probability of returning to the hegemon state is sufficiently high. This is because the only benefit that a dominant country might have of retaining $R$ is if returns to the hegemon state (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 (from Proposition 8 we know that $R$ will not be maintained in the bipolar state).

To state this result precisely, let a symmetric transition matrix $Q$ be defined as a transition matrix in which $q_{h_i,d_i} = q_{h_j,d_j}$, $q_{d_i,h_i} = q_{d_j,h_j}$, $q_{d_i,\Phi_i} = q_{d_j,\Phi_j}$ and $q_{\Phi_i,d_i} = q_{\Phi_j,d_j}$, $\forall i, j$ with $i \neq j$. Moreover, say that $Q$ is "smooth" if $q_{d_i,h_j} = q_{d_i,d_j} = q_{h_i,d_j} = q_{h_i,\Phi} = q_{\Phi,h_i} = 0, \forall i, j$ with $i \neq j$.

Proposition 8 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, it may be possible that an equilibrium in which hegemons and dominants choose rules on-equilibrium path exists (but, of course, not on the bipolar state). We prove the following necessary condition for such equilibrium.

**Proposition 9 (Hegemonic Stability Theory for Bipolar Worlds)** *In a bipolar world under assumption A2, if $Q$ is symmetric and smooth, a pure strategy symmetric Markovian equilibrium where hegemon countries create and maintain $R$ and dominant countries sustain $R$ exists only if the probability of a dominant country returning to hegemon is sufficiently high.*

# 6 A Leaderless 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 leaderless scenario, that is, a state in which the world lacks a clear leader. We study the feasibility of a rules-based trade regime in such a world.

Specifically, suppose that there is an extra state in $\Omega$ where there is neither a dominant nor a hegemon country. Instead, in this extra state, all countries are symmetric. Therefore, we now have an extended state space, denoted by $\tilde{\Omega} = \Omega \cup \{\mathcal{N}\}$, where $\mathcal{N}$ represents this extra state where no country is either hegemon or dominant. This state represents a situation where the world lacks a clear leader, and the choice of rules- versus power-based regimes depends on the actions of all $N$ countries. 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.

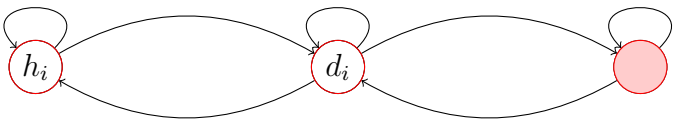



Figure 12: Transition between states; leaderless world

Whenever the system reaches state $\mathcal{N}$, all $N$ countries play a simultaneous regime-choice game. This game has the following properties. We assume that unanimity 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: Leaderless 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 supported $R$ also obtains the payoff corresponding to $P$ while still paying $c'$ for supporting the switch.[23] 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: Leaderless 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 leaderless world, $c'$ aside, all countries prefer to be in a rules-based regime. Although the matrices above present only the static payoffs, when considering the equilibrium continuation values that we 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.[24]

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 $\sigma(\mathcal{N}, R) = P$;

[23] While the exact assumptions of how rules might be sustained and implemented in the leaderless state are somewhat ad hoc, the qualitative message of this section is still valid if we relax 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 and consider the case in which $R$ is implemented in state $\mathcal{N}$ if a subset of countries strictly smaller than $N$ chooses $R$.

[24] 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 will keep it in equilibrium.

(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 leaderless world. Under such 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 leaderless state. We formalize this result below.

**Lemma 2 (Maintaining Status Quo at Leaderless State)** *Consider a world that has a leaderless state. If there exists a rules-based symmetric Markovian equilibrium in which 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 leaderless 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 leaderless state than in a world that may transition to a bipolar state.

**Proposition 10 (Leaderless Admits More Rules-Based Equilibria than Bipolar Systems)** *Consider a world with a leaderless state and another world with a bipolar state. If both worlds have the same transition matrix $Q$ and 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 leaderless state, then there are no rules-based equilibria when the world has a bipolar state.*

We establish this by showing that when countries adhere to the status quo in the leaderless 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. Thus, it is easier for many small countries without market power to coordinate on $R$ than for a few countries with market power to do so.

# 7 Conclusion

In the last decade, the rules-based multilateral trading system has shown alarming signs of instability, with countries displaying increasing disregard for established rules and growing reliance on economic power in trade relations. This is worrying because that system is believed to have paved the way for significant gains from trade since the late 1940s. Moreover, restoring it in the future may be difficult once it has been dismantled.

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 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. Our theoretical framework, with a dynamic stochastic game and focus on symmetric Markovian equilibria, allows us to analyze the equilibria in these long-term interactions and to study 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. Furthermore, our findings 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 the cost 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 definition of the trade regime, additional challenges arise. In particular, we show that incentives to exploit market power in international markets and incentives 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 – for example when power asymmetries are sufficiently small. We find that, because market power forces are weaker in such a leaderless state, it is easier to sustain a rules-based equilibrium in that case than in an environment in which there is a bipolar state.

Our framework, when viewed more broadly, helps explain recent shifts in multilateral trade cooperation. For decades following the 1940s, the U.S. was a clear hegemon, upholding a rules-based regime. However, its dominance has steadily eroded over time. In the context of our model,

the U.S. can now be seen 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. Our analysis suggests that, in such a state, market power and free-riding effects make it 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 we may already be in one, and for that reason facing immense difficulties to sustain the inherited rules-based system.

Is there hope? Our model, in its current form, cannot fully answer this question because it takes the existing set of "rules" – which we interpret as those established under the GATT/WTO framework – as given. In that sense, it is inherently *fatalist*: it helps us understand when a rules-based equilibrium is sustainable but offers no path for reform when it is not. In reality, as the global economy evolves, so too must the rules governing it. This calls for considering alternative frameworks that not only address modern challenges but are also politically and economically feasible. Indeed, recent proposals by several scholars have explored new frameworks for WTO rules and disciplines (Grossman and Sykes, 2025; Mavroidis, 2025; Rodrik, 2024; Stiglitz and Rodrik, 2024). A recurring theme in these proposals is that a more flexible system – let's call it "rules-lite" – could improve long-term viability.

This is broadly consistent with the mechanism of our model. The model stresses that the key reason why a rules-based system may fail is because dominant countries obtain a larger static payoff from *power* than from *rules*. If the rules are redesigned to narrow this payoff gap, they will make dominant countries more willing to tolerate *rules-lite*. However, if the current system is globally efficient, such a redesign will also inevitably reduce the benefits that smaller, non-leading countries receive.

At first glance, this may seem irrelevant to the system's sustainability, since non-leading countries are not pivotal in determining the regime. But from a dynamic perspective, the opposite is true: dominant countries only accept short-term losses under a rules-based system because they anticipate future gains when they eventually lose dominance. If these future gains are lower under *rules-lite*, then the long-run appeal of a rules-based system for dominant countries also weakens.

This discussion highlights a fundamental tradeoff in WTO reform: any attempt to enhance the system's sustainability must balance the concessions made to dominant countries against the losses imposed on smaller economies. Future research should explore this tradeoff further to design the most viable system of rules.

# 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 their 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 aggregate national income and its rate of growth. 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 benefits 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.[25] 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 the author 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.

As 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 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 probabilis-

[25]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.

tic 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 may 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 remains the same as in the aggregate data (Figure 1), with protectionist measures consistently surpassing liberalizing ones. 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, reflecting a broader global trend.

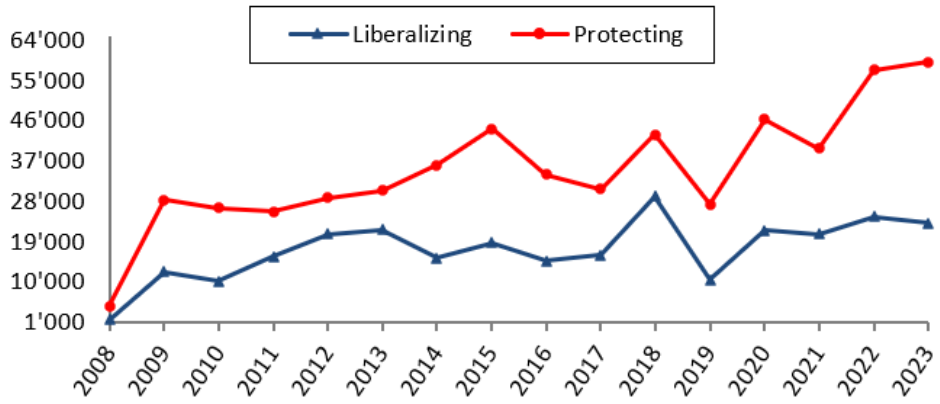



Figure 13: World trade policy interventions, excluding U.S. and China (2008 - 2023)
Source: Global Trade Alert database, 2024

# C Static Payoff Ordering

Figure 14 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.[26] If the hegemon chooses a power-based regime, because of the asymmetry of power, the gains from trade (GFT) would be 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. The hegemon's best alternative is therefore to propose a rules-based system. In that case, the rules would neutralize the effect of power asymmetry in negotiations and the GFT are split more evenly. This outcome is represented by point $R$.[27] 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.

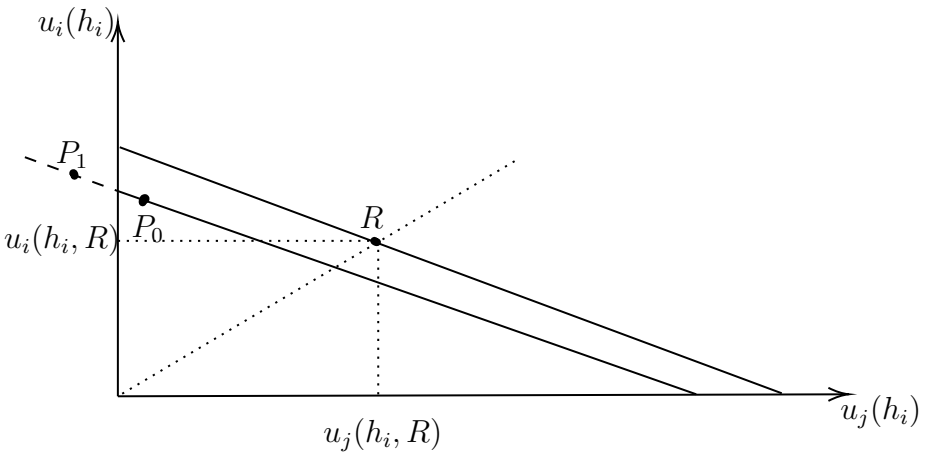



Figure 14: Stage game payoffs - Hegemon case

Figure 15 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 subordinate country is still better off than without any trade agreement. Therefore, the dominant country does not need to support a rules-based system to induce participation.[28]

[26]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.

[27]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.

[28]Observe that, as drawn, $u_i(d_i, P)$ in Figure 15 may look higher than $u_i(h_i, R)$ in Figure 14. But even if

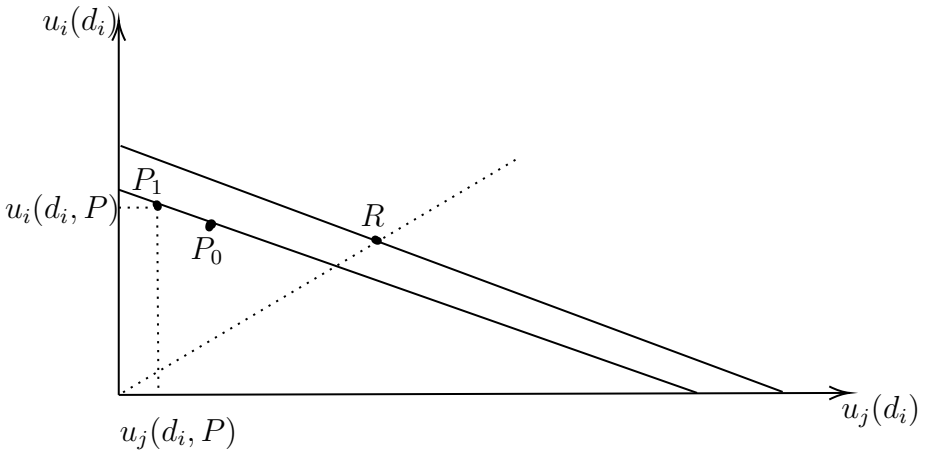



Figure 15: Stage game payoffs - Dominant case

$u_i(d_i, P) > u_i(h_i, R)$, a hegemon may still not want to give up its hegemony to achieve a higher payoff. 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.

# D Mathematical Appendix

**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 $(\omega, \rho)$, denoted by $\tilde{V}(\omega, \rho)$. Before we proceed, note that this strategy $\sigma$ 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:

$$u_i(h_i, R) + \delta \sum_{\omega'} q_{\omega,\omega'} \tilde{V}(\omega', R) - c \geq u_i(h_i, P) + \delta \sum_{\omega'} q_{\omega,\omega'} \tilde{V}(\omega', P). \tag{4}$$

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

$$\forall c < \frac{u_i(d_i, P) - u_i(d_i, R)}{\delta q_{d_i,h_i}}, \tag{5}$$

we have that:

$$u_i(d_i, P) + \delta \sum_{\omega'} q_{d_i,\omega'} \tilde{V}(\omega', P) \geq u_i(d_i, R) + \delta \sum_{\omega'} q_{\omega,\omega'} \tilde{V}(\omega', R). \tag{6}$$

Conditions (4) and (6) imply that the proposed strategy profile $\sigma$ induces a symmetric Markovian equilibrium; indeed, it induces the cyclic equilibrium. Moreover, if we decrease $\delta$, condition (5) still holds and so do conditions (4) and (6). 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:

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

This can be simplified to:

$$c > u(\omega, R) - u(\omega, P) + \delta \sum q_{\omega,\omega'}(\tilde{V}(\omega', R) - \tilde{V}(\omega', P)). \tag{7}$$

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 (7) is satisfied and thus there is a power-based equilibrium. We again defer the proof of uniqueness to Lemma 3. ■

**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,

$$u_i(d_i, R) + \delta \sum_{\omega' \in \Omega} q_{d_i,\omega'} V_i(\omega', R) \geq u_i(d_i, P) + \delta \sum_{\omega' \in \Omega} q_{d_i,\omega'} V_i(\omega', P), \tag{8}$$

that is, $V_i(d_i, R) \geq V_i(d_i, P)$. We also know from Proposition 3 that if $\sigma$ 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

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

that is, $V_i(d_i, P) > V_i(d_i, R) - c$. Combining (8) and (9) gives us the following inequalities:

$$c > \Delta V_i(d_i) \geq 0, \forall i. \tag{10}$$

For $c = 0$, inequality (10) cannot hold. By continuity of the value functions $V$ and the fact that $V_i(d_i, R)$ is not affected by $c$, 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. The 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:

$$\begin{aligned} V_i(d_i, R) &= u_i(d_i, R) + \delta \sum_{\omega} q_{d_i,\omega} V_i(\omega, R) \\ &\leq u_i(d_i, R) + \frac{\delta}{1-\delta}\bar{u}, \end{aligned} \tag{11}$$

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

$$\begin{aligned} V_i(d_i, R) &\geq u_i(d_i, P) + \delta \sum_{\omega} q_{d_i,\omega} V_j(\omega, P) \\ &\geq u_i(d_i, P) + \frac{\delta}{1-\delta}\underline{u}, \end{aligned} \tag{12}$$

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 $\delta$ and (ii) $\Delta \equiv \frac{\delta}{1-\delta}(\bar{u} - \underline{u})$ is increasing in $\delta$. Since $u_i(d_i, P) > u_i(d_i, R)$, there exists $\underline{\delta} > 0$ such that

$$u_i(d_i, P) > u_i(d_i, R) + \frac{\delta}{1-\delta}(\bar{u} - \underline{u}), \ \forall \delta < \underline{\delta}. \tag{13}$$

Therefore, for such $\delta < \underline{\delta}$, (13) contradicts (11) and (12). ■

**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 $\omega$ 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

$$u_i(h_i, R) + \delta \sum_{\omega' \in \Omega} q_{h_i,\omega'} V(\omega', R) \geq u_i(h_i, P) + \delta \sum_{\omega' \in \Omega} q_{h_i,\omega'} V(\omega', P), \tag{14}$$

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 $\sigma^*$ is a symmetric equilibrium strategy profile with $\sigma(d_i, P) = R$ for some $i \in \{1, 2, ..., N\}$. Since $\sigma^*$ 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 $\sigma'$ and show that it improves upon the proposed strategy $\sigma^*$, implying that $\sigma^*$ is not an equilibrium, a contradiction.

Let us construct this deviation strategy $\sigma'$ in a way that it is identical to $\sigma^*$ at all histories except (at most) two: (i) at a history $h'$ in which the pair $(\omega, \rho)$ 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 $\sigma^*$ specifies the play of $R$ at such history), whereas at history $h''$ the constructed deviation strategy $\sigma'$ specifies the play of $R$ (which may or may not coincide with the action specified by $\sigma^*$ at this history $h''$).

Consider country $i$'s payoff having reached history $h'$ and following the equilibrium strategy $\sigma^*$. 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 $\sigma'$ 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 $\sigma'$ is better than the payoff following the proposed equilibrium strategy $\sigma^*$.

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 $\sigma^*$, 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 $\sigma^*$ the cost is paid immediately, whereas under $\sigma'$ 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 $\sigma'$ at this history is identical to $\sigma^*$ and it specifies that country $j$ will choose $R$, i.e. $\sigma^*(d_j, P) = R$. Moreover, under the equilibrium strategy $\sigma^*$, 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 $\sigma'$ implies a higher current payoff for country $i$ at history $h'$ than the proposed equilibrium strategy $\sigma^*$ and an identical continuation payoff. Thus, $\sigma^*$ cannot be a subgame perfect Nash equilibrium. This is a contradiction. ■

**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 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 $\delta$ and any given equilibrium, there exists $\Delta v$ such that $V(\omega, R) - V(\omega, P) \leq \frac{\Delta v}{1-\delta}, \forall \omega$. Moreover, there exists $\underline{v}$ such that $V(\omega, P) \geq \frac{\underline{v}}{1-\delta}, \forall \omega$, and $\bar{v}$ such that $V(\omega, R) \leq \frac{\bar{v}}{1-\delta}, \forall \omega$.[29] Then, for any equilibrium, we can

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

write the expected continuation payoff at $(d_i, P)$ as:

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

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

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

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

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

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

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

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:

$$u_i(d_i, P) - u_i(d_i, R) \leq \delta \frac{q_{d_i,d_i}}{1 - \delta q_{d_i,d_i}} \{ u_i(d_i, R) - u_i(d_i, P) + \delta (1 - q_{d_i,d_i}) \frac{\bar{v} - \underline{v}}{1 - \delta} \} + \delta (1 - q_{d_i,d_i}) \frac{\Delta v}{1 - \delta} \tag{15}$$

As $q_{d_i,d_i} \to 1$, we can see that the r.h.s. of (15) 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 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 $(\delta, c)$ 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 $(\delta, c)$ 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 $(\delta, c)$ 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 $\omega$ at period $t$.

$$c \leq \sum_{t\geq 1}\sum_{i\in\mathcal{I}} \delta^{t-1}\{u_i(\omega, R) - u_i(\omega, P)\}\pi(\omega, t). \tag{16}$$

Given our assumption above, we have that $u_i(\omega, R) - u_i(\omega, P) > 0$ for all $\omega \in \Omega$. Thus, the right-hand side of the inequality (16) is decreasing in $\delta$. Additionally, note that the set $S$ of parameters $(\delta, c)$ 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 $(\delta, c)$ with high $\delta$. Formally, $\forall c$, there exists $\delta_c$ 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 $\delta$ 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 $\delta$, a sufficiently high cost exists such that the efficient solution prescribes the constant choice of a power-based regime. Suppose not. Then,

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

where $\underline{u_j}(P)$ represents each player $j's$ minimum payoff at any state $\omega$. 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 $(\delta_1, c_1)$ and $(\delta_2, c_2)$ under which the optimal solutions species never paying the switching cost $c$. Then, following (16), we can write:

$$c_k \geq \sum_{t\geq 1}\sum_{i\in\mathcal{I}} \delta_k^{t-1}\{u_i(\omega, R) - u_i(\omega, P)\}\pi(\omega, t), \quad k = 1, 2. \tag{17}$$

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

$$\lambda c_1 + (1-\lambda)c_2 \geq \sum_{t\geq 1}\sum_{i\in\mathcal{I}}(\lambda\delta_1 + (1-\lambda)\delta_2)^{t-1}\{u_i(\omega, R) - u_i(\omega, P)\}\pi(\omega, t), \quad k = 1, 2, \tag{18}$$

which concludes our proof. ■

**Proof of Proposition 6.** The proof of this proposition is by construction. We use the same parameters as in our benchmark simulations but we add 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 results:

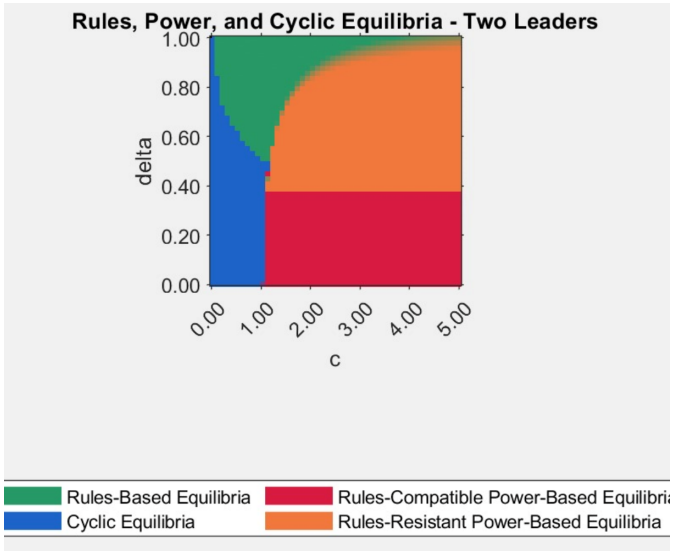



Figure 16: Rules-Based Equilibria in Bipolar World

■

**Proof of Proposition 7.** 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

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

At the same time,

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

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

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

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

**Proof of Proposition 8.** 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 9.** 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:

$$u_i(d_i, R) + \delta \sum_{\omega'} q_{d_i,\omega'} V(\omega', R) \geq u_i(d_i, P) + \delta \sum_{\omega'} q_{d_i,\omega'} V(\omega', P) \tag{19}$$

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

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

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

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 $\Phi$ (Proposition 8), we have that $V(\Phi, R) = V(\Phi, P)$. Then, we can rewrite the condition above as:

$$V(d_i, R) = \frac{u_i(d_i, R) + \delta(q_{d_i,h_i}(V(h_i, P) + c) + q_{d_i,\Phi} V(\Phi, P))}{1 - \delta q_{d_i,d_i}}. \tag{22}$$

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:

$$u_i(d_i, P) - u_i(d_i, R) \leq \delta q_{d_i,h_i} c. \tag{23}$$

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

$$u_i(d_i, R) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', R) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, P)) \geq u_i(d_i, P) + \delta(\sum_{\omega' \in \Omega \setminus \mathcal{N}} q_{d_i,\omega'} V_i(\omega', P) + q_{d_i,\mathcal{N}} V_i(\mathcal{N}, P)). \tag{24}$$

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

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

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

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

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

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

Using a reasoning analogous to Lemma 1, it cannot be true that (24), (25) or (26) hold but (27) does not. ■

**Proof of Proposition 10.** Suppose that there exists a rules-based equilibrium in a bipolar world. This implies that

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

This condition implies that in the leaderless world under the strategy in which countries maintain the status quo at the leaderless 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 leaderless world). Thus, if there is a rules-based equilibrium under bipolar, there must be at least one rules-based equilibrium under a leaderless state. ■

# 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.

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. *Mimeo*, 2024.

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

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

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

G. Grossman and A. Sykes. International discipline of government subsidies: An assessment of current rules and possible reforms. *Mimeo*, 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.

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

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

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. doi: 10.1257/aer.97.4.1374.

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*, 2024.

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.

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

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.

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. Hegemonic globalization. *Mimeo*, 2024.

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

World Bank. World development indicators, 2025. URL https://databank.worldbank.org/.

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