International Coordination and the Informational Rationale for Delegation*

Emiel Awad[†]
London School of Economics and Political Science
Department of Government

Nicolás Riquelme[‡]
Universidad de los Andes, Chile
School of Business and Economics

June 13, 2022

Abstract

In contrast to existing rationales for delegation that are centered on international organizations' (IO) expertise, we study delegation if countries have information that IOs lack. In our model, countries transmit information through cheap talk and costly signals and benefit from international coordination and adapting policies to domestic circumstances. Even though countries gain due to an IO's ability to help with coordination, they waste more resources in signaling once the involvement of IOs induces more coordinated policies. With too much uncertainty, countries find it optimal to coordinate outside IOs. Additionally, more disagreement between countries strengthens their willingness to delegate but makes enforcement of cooperative agreements more difficult. Further, IOs may be designed with limited discretion to reduce the scope for costly signaling. Although international relations theories generally state that IOs help with information transmission, our results further our understanding of how IOs may negatively affect countries' incentives to share information.

*We thank Jason Davis, Zuheir Desai, Torun Dewan, Stefan Elbl, Mark Fey, Jacque Gao, Michael Gibilisco, Gleason Judd, Rafael Hortala-Vallve, Vanand Meliksetian, Jack Paine, Bradley Smith, Randy Stone, Stephane Wolton, several anonymous reviewers, and audiences at the University of Rochester, London School of Economics and Political Science, and Quebec Political Economy Webinar Series for helpful comments. Riquelme gratefully acknowledges financial support from ANID/CONICYT (FONDECYT Iniciación #11220584).

[†]Postdoctoral Fellow, e.awad@lse.ac.uk.

[‡]Assistant Professor, nriquelme@uandes.cl.

There is extensive scholarly debate about when and why countries delegate to IOs.[1] The extant literature generally points to the benefits of delegation to international institutions. IOs are thought of as agents that solve international collective action problems, allow for credible commitments, and reduce transaction costs in policy-making (Abbott and Snidal, 1998; Hawkins et al., 2006; Bradley and Kelley, 2008; Koremenos, 2008; Hooghe and Marks, 2015). The literature identifies another benefit of IOs with incomplete information, generating an *informational rationale* for delegation. On one hand, if IOs have more information than countries, delegation leads to more informed policies. On the other, if countries have more information than IOs, then IOs help countries pool and share information (Koremenos, Lipson and Snidal, 2001, p. 788). In our paper, the focus is on the latter case where IOs are thought to help with international cooperation even though countries are better informed than IOs.

A sizable game theoretic literature covers the case where IOs have an informational advantage over countries. For example, Johns (2007) studies the selection of UN bureaucrats who provide information to bargaining member countries. Fang and Stone (2012) analyze how IOs provide policy recommendations to member countries, applied to the WHO and IMF. Crombez, Huysmans and Van Gestel (2017) study the European Commission and their role as an agenda setter that provides information to the European Parliament and Council.[2] Such papers build on a vast principal-agent literature in which a principal may give an agent decision-making authority if the agent has more expertise than the principal and preferences are sufficiently aligned.[3]

In the other part of the informational rationale for delegation, IOs are theorized to help countries share information and reduce uncertainty. For example, Keohane (1984) notes that "By reducing asymmetries of information through a process of upgrading the general level of available information, international regimes reduce uncertainty (p. 94)." Similarly, Thompson (2015) highlights the informational benefits of IOs in monitoring and

[1]See Voeten (2019) for a review of this debate.

[2]See also Chapman (2007) who studies how leaders choose which institutions to consult and Fang (2008) who studies an international institution that may give recommendations to national governments on foreign policy issues.

[3]See, e.g., Bendor, Glazer and Hammond (2001) for a review.

exchanging information: "Moreover, the exchange of information and discourse that takes place within an IO tends to reveal information about countries preferences and intended actions [...], leading to more effective monitoring and higher quality signaling at the international level (p. 30)."[4] These informational benefits of IOs can be quite significant, given that incomplete information may be a cause of international conflict (Fearon, 1995) or cause additional issues in collective action problems (Kenkel, 2019).

It is, however, an open question whether IOs *always* solve informational issues among countries even if countries have an informational advantage over an IO.[5] That is, what are the scope conditions for the informational rationale for delegation to hold? It is relatively obvious that IOs may help if they have specific expertise that countries do not have or if they can monitor countries and acquire information about them and aggregate it.[6] What happens, though, if countries have more information than the IO and the IO has no way to access this private information directly—does the IO provide informational benefits for countries to pool and share information? It is unclear why formalized institutions are necessary or beneficial to help countries transmit information compared to *ad hoc* cooperation (Keohane, 1982, 1984; Morrow, 1994; Koremenos, Lipson and Snidal, 2001).

Thus, to make progress on this question, we develop a formal model with incomplete information, in which countries both value international coordination and adaptation to the domestic political and economic environment.[7] Two countries have private information

[4]Also, Abbott and Snidal (1998): "They keep track of agreements on particular issues, trade-offs, and areas of disagreement, periodically producing texts that consolidate the current state of play. They also transmit private offers or assurances, improving the flowing of information. (p. 12)" Also, Barnett and Finnemore (1999, 709), however, also hint at potential negative consequences of IOs referring to informational issues: "As IOs create transparencies and level information asymmetries among countries (a common policy prescription of neoliberals) they create new information asymmetries between IOs and countries. Given the neoliberal assumption that IOs have no goals independent of countries, such asymmetries are unimportant; but if IOs have autonomous values and behavioral predispositions, then such asymmetries may be highly consequential."

[5]There is relatively little focus on situations in which countries are better informed than IOs but see, however, Carson and Thompson (2014) and Carnegie and Carson (2019).

[6]See, e.g., McAllister and Schnakenberg (2021) for a formal model in which IOs acquire information about countries for international climate agreements.

[7]Our setup is similar to Alonso, Dessein and Matouschek (2008) and Rantakari (2008), who study an organization with an uninformed headquarters and multiple divisions that are better informed about local conditions, where the headquarters cares about coordination and adaptation. The main innovations of our model are that we consider costly signals in addition to cheap talk and that our modeling

about their domestic circumstances and can transmit their information through cheap talk and costly signals to each other or the IO. Apart from communicating their preferences in written or spoken form, countries may engage in activities, such as organizing conferences and subsidizing firms, to potentially signal to other countries how strongly they prefer specific policies.[8] Once information is shared, policies are made depending on the institution.

We compare signaling and policy-making without and with IOs. Without IOs, two countries make policies absent external influence. If there is an IO, however, it makes policies on the two countries' behalf. We emphasize strategic considerations and downplay exogenous costs or benefits of information transmission and coordination in IOs. That is, our results do not rely on the fact that IOs are costly to create or that countries cannot transmit information without IOs. We are interested in the strategic effects of institutions on international cooperation and information sharing.

Our model produces two sets of results, relating to the effect of IOs on coordination and signaling. First, on the plus side, international delegation always increases coordination. In isolation, countries take decisions without internalizing the effects on other countries and coordinate too little compared to what they could have achieved if they had the ability to commit themselves to certain promised decisions. The influence of IOs allows countries to lock in more coordination than they could achieve individually and they benefit from this. This is a well known result in the literature on the rational design of IOs (Koremenos, Lipson and Snidal, 2001).

The second result is novel. It states that international delegation and the involvement of IOs leads to more costly signaling. We show that international institutions do not help countries to share more information; countries share the same amount of information both in the presence and absence of an IO. The costliness of information transmission, however, is always greater when policymaking occurs in the presence of international

framework delves deeper in different directions with respect to institutional design, such as choosing the IO's preferences and limiting its discretion.

[8]Note, these activities may be productive in their own right, but countries may spend more on these activities than is individually optimal as a costly signal that is observed by other countries.

institutions. One reason is that following delegation, countries have to incur costs to influence a decision that they would have made themselves absent delegation. Even if the IO is less biased than the other country, the negative effect of the IO's increase in authority always outweighs the positive effect of transmitting information to a less biased receiver (Crawford and Sobel, 1982; Austen-Smith and Banks, 2000). In addition, compared to each country in isolation, the IO cares more about coordination than adaptation to domestic circumstances. As a result, IOs more quickly adapt policies to countries' circumstances as long as these policies are coordinated. This leads to stronger incentives to misrepresent by countries, increasing the costliness of signals to transmit information. This happens because policies are more sensitive to what information countries transmit, making lying more attractive.

Taken together, these two results imply that it is not immediate that countries benefit from IOs if they want to coordinate and adapt to domestic conditions under incomplete information. Even if coordination is crucial to countries, too strong incentives to send costly signals in IOs may make countries worse off. We find that the problem of increased costly signaling is more pronounced with greater uncertainty. Thus, for delegation to be beneficial, there must not be too much uncertainty to ensure that the benefits from increased coordination outweigh the costs of deteriorated signaling.

We extend our model to ensure robustness to different decision-making protocols and parametric assumptions. First, we show that it is equivalent to let countries bargain rather than allowing an agent to make decisions. Thus, it is not necessary for our results that IOs make decisions on the countries' behalf. The second extension emphasizes, however, that a key function of IOs is to enforce its decisions. Our results only hold if the IO has sufficient enforcement capabilities, while there is less coordination and money burning with relatively weak IOs. Further, more disagreement between countries increases their willingness to delegate but makes enforcement of more coordinated decisions difficult. A fourth extension analyzes an environment where only one country's preferences are uncertain and shows that incentives to delegate are stronger. Fifth, if two countries care differently about coordination, such as a large and small country, more authority should be given to the larger country that cares less about coordination, to dampen money burning incentives in the presence of more severe uncertainty. Finally, we study an environment in which international bureaucrats may make policy with limited discretion. Although this decreases how adapted decisions are to countries' domestic circumstances, it brings about less wasteful money burning.

The applicability of our model is subject to several scope conditions. First, the IO has to influence what policies countries make through its enforcement capabilities. Our model does not apply when IOs only serve as a way to communicate without any sort of enforcement mechanism. Indeed, in an extension we show that our results are weaker if the IO has fewer enforcement capabilities. Second, we study situations in which countries have an informational advantage over the IO (Stone, 2009, 35–36). This setting is especially relevant for, e.g., newly formed IOs, IOs that operate with limited budgets, or if countries have access to private information of domestic firms and markets that IOs cannot obtain or easily verify. Third, we study policy domains in which countries want to coordinate policies. They do not merely prefer to get their ideal point for their own country but they also are affected by other countries if policies are dissimilar. This is important because it generates incentives to misrepresent information; there clearly are no benefits of IOs if countries always prefer to tell the truth regardless.

There are multiple policy areas that fit our model to various degrees. One example comes from the EU, where member states recently agreed to climate neutrality by 2050.[9] First, tackling climate change is a collective action problem in which multiple countries need to cooperate and exert enough effort regulating their economies. Another aspect, however, is a coordination problem among countries. At the time of negotiations, it was a contentious issue whether nuclear power may be part of a member state's strategy to reduce greenhouse gas emissions. The European Commission offered two possible policy options to its member states, where nuclear power would either get a *green label* or not. This green label would have an impact, as although each country potentially has an individual benefit of using nuclear power, it would also benefit if other countries would use the same

[9] `https://ec.europa.eu/clima/policies/strategies/2050_en` (accessed on April 20, 2021).

energy sources. This is because it could spur technological advancement and yield cheaper energy production due to economies of scale. That is, France has an individual incentive to push for a green label to help its domestic nuclear energy producers, and there are potentially further benefits if other countries coordinate on the same technology. On the other hand, countries such as Germany are opposed to nuclear energy, referring to it as a dangerous energy source, and plan to use hydrogen energy sources among others. This brief case highlights how countries may have different preferences based on their domestic economy and value coordination on specific policies.

The model can also be applied to organizations other than IOs. The common feature is between centralization and decentralization with incomplete information about *local conditions*. For example, in firms with a headquarters and multiple divisions that deal with different products or regions, a trade off exists between coordination and adaptation, and the division managers know more about the technical aspects than the headquarters (Alonso, Dessein and Matouschek, 2008; Rantakari, 2008). A similar trade-off could also apply to joint ventures between firms because they can create a new unit and must decide how much authority to give to the centralized unit.[10] Alternatively, our model could also apply to situations in which multiple states operate in the presence or absence of a federal government. States know better than the federal government what policies are optimal for their respective populations, but may gain from inter-state coordination. As a result, a federal government may induce more coordination but the private information of individual states still needs to be aggregated. In the remainder of our paper, however, we focus our attention on IOs and the specific trade-offs in the international domain.

The Model

Consider two countries that want to coordinate their decisions but disagree on *which* decisions and are incompletely informed about the other country's preferences. We compare two institutions that symbolize the absence and presence of an IO. In the first, countries

[10] An extreme form of centralization could be a merger between firms, in which the individual units no longer exist.

act without a formalized structure and make their own decisions. In the second, countries delegate to an IO that makes decisions on their behalf. We emphasize that this second version is assumed for the sake of exposition, and we provide a formal justification in an extension of the model.

The timing is as follows. First, Nature draws two random variables, each one being private information for each country. These types determine a country's bliss point. The type of 1 (θ_1) is drawn from a uniform distribution with support $\Theta_1 = [\underline{\theta}_1, \overline{\theta}_1] = [-1-s, -1+s]$. Similarly, 2's type θ_2 is independently drawn from a uniform distribution with support $\Theta_2 = [\underline{\theta}_2, \overline{\theta}_2] = [1-s, 1+s]$. Note that country 1's type distribution has a mean of -1, while country 2's distribution has mean 1. The value $s \in [0, 2)$ indicates how much uncertainty there is about each actor's private information.[11] In the second stage, each country $i \in \{1, 2\}$ observes its private information θ_i and sends a signal (b_i, m_i), where $b_i \geq 0$ is the amount of burned money, which enters negatively in country i's payoff, and m_i is a cheap talk message.

The two institutions differ in the third stage. In the first model (the *no-delegation game*) each country takes a decision while in the second model (the *delegation game*), both countries' decisions are delegated to the IO. In the third stage of the no-delegation game, each country i observes how much the other country has burned (b_j) and what cheap talk message was sent (m_j), and takes decision $d_i \in \mathbb{R}$. In the delegation game, authority is delegated to an agent, conceptualized as an IO, who takes both decisions on the behalf of countries (the two principals). The IO observes amounts of burned money (b_1, b_2) and messages (m_1, m_2), and takes decisions $(d_1, d_2) \in \mathbb{R} \times \mathbb{R}$.

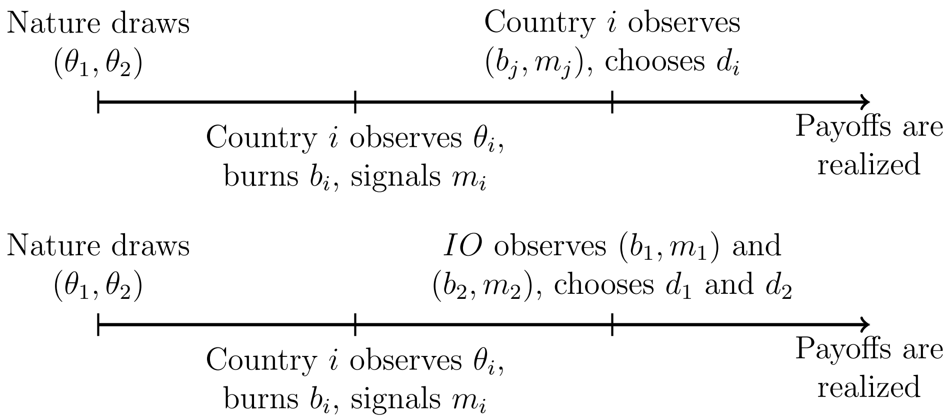
Each country's utility consists of a policy payoff and the cost of burned money. This utility can be written as $u_i(d_i, d_j, \theta_i, b_i) = \pi_i(d_i, d_j, \theta_i) - b_i$, where its policy payoff is

$$\pi_i(d_i, d_j, \theta_i) = -(1-\beta)(d_i - \theta_i)^2 - \beta(d_i - d_j)^2.$$

This policy payoff consists of two parts. The first term is a country's adaptation motive,

[11] The results are similar when $s \geq 2$, but it complicates exposition; we discuss this in the online appendix.

Figure 1: Timing without and with Delegation



Note: The top panel illustrates the no-delegation game and the bottom panel shows the delegation game.

where $-(d_i - \theta_i)^2$ is the cost of decisions that are not adapted to domestic conditions. The further d_i is from a country's bliss point θ_i, the higher is the cost that country i incurs. The second term measures coordination, where $-(d_i - d_j)^2$ is the cost of uncoordinated decisions.[12] The parameter $\beta \in (0, 1)$ measures the importance of coordination relative to tailoring decisions to domestic conditions. All in all, country 1 is best off if $d_1 = d_2 = \theta_1$ and both countries coordinate on 1's bliss point θ_1. Symmetrically, country 2 is best off if both coordinate on 2's bliss point with $d_1 = d_2 = \theta_2$.

Given delegation, we study a symmetric setup and assume that the IO weighs both countries' policy payoff equally.

$$u_{IO}(d_1, d_2, \theta_1, \theta_2) = \frac{1}{2}\left[\pi_1(d_1, d_2, \theta_1) + \pi_2(d_1, d_2, \theta_2)\right].$$

As we analyze two models, equilibrium definitions depend on the set of strategies that every actor has. In the no-delegation game, country i's strategy is (i) a mapping from types θ_i to signals (b_i, m_i) and (ii) from types θ_i, country i's and j's signals to decisions d_i. In the delegation game, a strategy of country i consists of a mapping from types to signals; and for the IO it is a mapping from the signals of both countries to decisions d_1 and d_2.

[12] We assume that the payoffs of both countries are equally concave, but the analysis goes through if they differ. See Cressman and Gallego (2009).

We study perfect Bayesian equilibria (PBE), where players use sequentially rational strategies and update beliefs using Bayes' rule wherever possible. A strategy is sequentially rational if it is a best response to other players' strategies at any point in time. The requirement of Bayes' rule means that players have beliefs that are consistent with other players' strategies. We fully specify our equilibrium concept in the appendix.

A general issue in signaling models with cheap talk and money burning is the existence of multiple equilibria.[13] Indeed, Lemma 0 in the Appendix uses existing results in the literature to characterize the set of equilibria, of which there are infinite. In a share of signaling models, refinements may delete a large set of these equilibria, potentially guaranteeing a unique one.[14] We adopt a commonly used refinement in models with costly signaling; the monotonic D1 refinement (mD1).[15] The mD1 refinement puts strong restrictions on beliefs off the path of play which ensures a unique equilibrium in our model.[16] We discuss this in more detail in the appendix. From now on, we call the unique PBE that survives the mD1 refinement an *equilibrium*.[17]

Lemma 1. *There exists a unique equilibrium in the no-delegation game and the delegation game. In this equilibrium, the amount that each country burns is fully informative about its type. It is without loss of generality to ignore cheap talk messages.*

Lemma 1 says that after observing how much money a country burned, every player perfectly infers the country's type on the equilibrium path. Thus, every equilibrium is outcome-equivalent regardless of the use of cheap talk messages. We therefore restrict

[13]Karamychev and Visser (2016) study equilibria that are optimal for the sender from an ex ante perspective. In Proposition 1, they show that existence of equilibria requires a partitional structure as in Austen-Smith and Banks (2000), and that any arbitrary partitioning of the type-space can be generated through the use of cheap talk and burned money. See also Kolotilin and Li (2021).

[14]Cho and Kreps (1987) and Chen, Kartik and Sobel (2008).

[15]Bernheim and Severinov (2003).

[16]D1 criteria require that the receiver does not attribute a deviation to a particular type if there is another type that is willing to make the deviation for a strictly larger set of possible receiver responses. The monotonicity assumption requires that money burning strategies are monotonic (weakly decreasing or increasing) in a country's type.

[17]It is important to note that this equilibrium does not maximize efficiency. See Karamychev and Visser (2016) for a discussion on *ex ante* socially optimal equilibria. Full separation is generally inefficient. Although the receiver's preferred equilibrium is fully separating, the sender prefers to reveal information through a finite number of intervals. Thus, full separation is efficient if and only if the receiver's payoff is sufficiently important relative to the sender's.

our attention to money burning strategies in describing equilibrium signaling strategies. In the appendix, after relaxing a parametric restriction on the amount of uncertainty s, the availability of cheap talk becomes necessary to characterize equilibria. Note that though the monotonicity requirement of mD1 implies that money burning strategies are *weakly* increasing or decreasing, in equilibrium, strategies are *strictly* increasing or decreasing.

To determine whether the involvement of IOs is beneficial, we are ultimately interested in countries' payoffs from an *ex ante* perspective. That is, we take an expectation over the payoff that countries obtain in equilibrium before types are drawn. From now on, in the expectation operator $\mathbb{E}[\cdot]$ we use the superscript 0 to denote prior beliefs and subscript i to designate player i's posterior beliefs. Our main comparison is between two institutions $\mathcal{I} \in \{ND, D\}$; the no-delegation game and delegation game respectively.

The uniqueness result on the set of equilibria implies that one advantage of our institutional comparison is that it is fair. This is because qualitatively similar equilibria are played given both institutions, and any player makes decisions based on full information about θ_1 and θ_2 on the equilibrium path. The differences between both institutions are characterized by which decisions are made and which signals are sent in equilibrium.

We decompose countries' welfare into two parts. One is *political* and purely measures policy choices without accounting for the informational aspect. Given an institution $\mathcal{I}$, there are two decisions $(d_1^{\mathcal{I}}(\theta), d_2^{\mathcal{I}}(\theta))$ which are ultimately a function of types $\theta = (\theta_1, \theta_2)$. Thus, we can write a country's policy payoff as $\pi_i^{\mathcal{I}}(\theta) := \pi_i^{\mathcal{I}}(d_i^{\mathcal{I}}(\theta), d_j^{\mathcal{I}}(\theta), \theta)$, and define a country's political welfare as follows.

Definition 1. Country i's *ex ante political payoff* in institution $\mathcal{I}$ is $\mathbb{E}^0[\pi_i^{\mathcal{I}}(\theta)]$.

We can perform a similar calculation for a country's *informational* welfare. Typically, this also measures the amount of information that is lost in signaling. Given Lemma 1, however, players are perfectly informed in equilibrium. The only welfare losses result from the amount of burned money, i.e., the expected costliness of the signal that a country sends. Under institution $\mathcal{I}$, country i burns $b_i^{\mathcal{I}}(\theta_i)$ given type θ_i. The following definition computes the expected costs of money burning. We define this as informational welfare because it wholly results from the presence of incomplete information.

Definition 2. Country i's *ex ante informational payoff* in institution $\mathcal{I}$ is $-\mathbb{E}^0[b_i^{\mathcal{I}}(\theta_i)]$.

Using Definitions 1 and 2, we can now precisely determine when a country is better off with an IO. Definition 3 states the condition precisely, highlighting how it depends on political and informational welfare.

Definition 3. Country i prefers to delegate if $\mathbb{E}^0[\pi_i^D(\theta)] - \mathbb{E}^0[b_i^D(\theta_i)] \geq \mathbb{E}^0[\pi_i^{ND}(\theta)] - \mathbb{E}^0[b_i^{ND}(\theta_i)]$.

Model Discussion

Before presenting the results, we explain the type of situation we aim to capture with our model. There is one decision that needs to be taken for each individual country. These countries have *bliss points* that encapsulate their domestic conditions and want to make decisions that are most suitable given these conditions. That is, they care about coordination and adaptation to domestic conditions. For example, when it comes to the setting of standards, countries may benefit from coordination due to network externalities in consumption and scale economies in production. At the same time, each country prefers the standard used by its firms. Each country must trade off these two incentives in policy-making as other countries do not find it optimal to be accommodating.

Information is private, as each country knows better what works best for itself than the other country does. Countries' information about their adaptation costs to common standards (Toulemonde, 2013), for example, is private. That is, each country knows in general what the other country wants, but not *how strongly* they want it. This information is transmitted through cheap talk and money burning (Crawford and Sobel, 1982; Austen-Smith and Banks, 2000). One interpretation of money burning is given by an example of TV standards in which countries attempted to influence their neighbors by assistance in marketing surveys, subsidies, exhibitions, fairs, and demonstrations (Crane, 1978). More generally, research and development, investments, or subsidies could also be partly interpreted as a form of money burning. The reason is that, although these may be productive activities, countries may spend more resources than is individually optimal to signal their resolve about the desire to coordinate on its preferred policy.

In the delegation game, the IO takes decisions on the countries' behalf. Although countries typically still retain policy authority when IOs are involved, we make this assumption to model an IO's member countries' bargaining in a reduced form. An extension with international bargaining confirms that this assumption is without loss of generality. Our results apply as long as IOs influence the costs and benefits of certain policy choices. That is, IOs must allow countries to obtain a certain degree of commitment to international policy choices. In addition, albeit stylized, the delegation game allows us to understand more realistic institutions with limited delegation through a comparison of two extreme institutions. IOs are not merely actors that provide policy recommendations to countries without enforcement, nor are they solely a forum to exchange information.

Results

The structure of our results section is as follows. First, we derive what decisions are made based on available information without and with an IO. Second, given decision-making strategies, we study how countries burn money to transmit information in equilibrium. Third, our main result takes these equilibrium strategies and computes the benefits of both countries without and with delegation.

Decision-making

Absent delegation, given Lemma 1, each country i is fully informed about all information θ_i and θ_j. In the final stage of the game, each country i takes a decision that weighs its adaptation and coordination motive. On the path of play, equilibrium decisions depend on the value of coordination β as follows. It places weight $1-\beta$ on its own type θ_i, which coincidentally is the weight of the adaptation motive. Further, country i places weight $\frac{\beta}{1+\beta}$ on the expected value of θ_j, which is driven by the need for coordination. Whatever country j signals about its type, country i wants to ensure its decision is more in line with

d_j. Finally, it places weight $\frac{\beta^2}{1+\beta}$ on the expectation that country j has about country i's type. The reason is that, similar to how country i responds to signaling by country j, the reverse is also true. Thus, by inducing a certain belief, country i has an expectation over what decision country j will take, and wants to coordinate with that decision.

Lemma 2. *In the equilibrium of the no-delegation game, country* $i = 1, 2$ *chooses*

$$d_i^{ND} = (1-\beta)\theta_i + \frac{\beta}{1+\beta}\mathbb{E}_i[\theta_j|b_j, m_j] + \frac{\beta^2}{1+\beta}\mathbb{E}_j[\theta_i|b_i, m_i].$$

There is a difference if the IO takes decisions d_1 and d_2. In this case, exclusively what matters is the IO's belief after observing both countries' signals. Recall that the IO weighs both countries' interests equally but does not take fully coordinated decisions, with $d_1 \neq d_2$. The reason is that it values adaptation to domestic conditions in both countries, which depends on θ_1 and θ_2. To summarize, though, the IO puts more weight on the benefits of coordination leading to more coordinated decisions.

Lemma 3. *In the equilibrium of the delegation game, the IO makes decisions* d_i^D *for countries* $i = 1, 2$

$$d_i^D = \frac{1+\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_i|b_i, m_i] + \frac{2\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_j|b_j, m_j].$$

The difference between Lemma 2 and 3 shows that delegation affects decision-making strategies. This means that decisions have different sensitivities to information, which ultimately affects how much money countries burn.

Signaling

Recall that Lemma 1 shows that countries completely separate using burned money, i.e., information is fully transmitted in equilibrium. Given that country 1 cares about 2's decision, 1 has an incentive to pull 2's decision d_2 closer to d_1. More specifically, 1 wants to pull d_2 downward, while 2 wants to pull d_1 upward. Thus, 1 has incentives to misrepresent its information and claim its type is lower than θ_1. The reverse is true for 2, who wants to say θ_2 is higher than it actually is. This prevents full information transmission through cheap talk and requires money burning.

How much must be burned by each type θ_i to ensure that all information is transmitted? This depends on how much each country is willing to lie about its information. More formally, two types θ_i and θ_i' must not find it profitable to mimic each other. As misrepresenting one's type may lead to a more preferred decision, money has to be burned to offset this incentive. The additional positive payoff after obtaining a more preferred decision must be exactly offset by the negative impact of burned money. Proposition 1 solves for this amount and summarizes equilibrium money burning strategies in the no-delegation model. Prior to presenting the results, it is instructive to define two functions of θ_1 and θ_2 that determine how money burning is affected by countries' information. Recall that the distribution of types for each country is defined as $\Theta_i = [\underline{\theta}_i, \overline{\theta}_i]$.

$$f_1(\theta_1) = \left(\theta_1 - \overline{\theta}_1\right)\left(\frac{\theta_1 + \overline{\theta}_1}{2} - 1\right)$$
$$f_2(\theta_2) = \left(\theta_2 - \underline{\theta}_2\right)\left(\frac{\theta_2 + \underline{\theta}_2}{2} + 1\right)$$

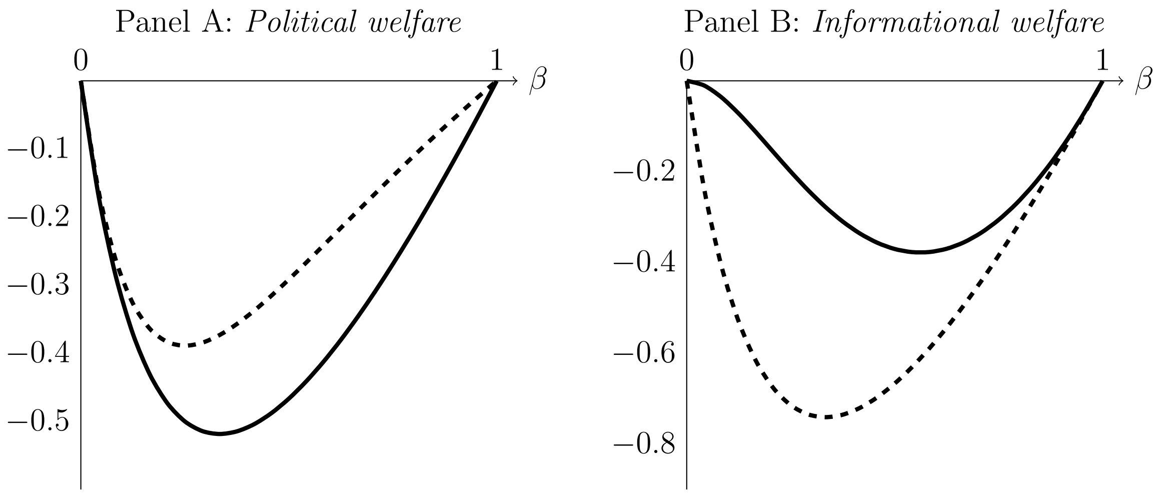
The function $f_1(\theta_1)$ increases when θ_1 moves *downward*, which means that country 1 burns more when its type is *lower*. Analogously, $f_2(\theta_2)$ increases when θ_2 moves *upward*, implying that country 2 burns more when its type is *higher*.

Proposition 1. *In the equilibrium of the no-delegation game, country i burns*

$$b_i^{ND} = \frac{2(1-\beta)\beta^2}{(1+\beta)^2} f_i(\theta_i).$$

The IO changes incentives to misrepresent information. Consider for instance country 1. Its information θ_1 still affects both decisions d_1 and d_2, but with delegation decisions are more coordinated regardless of adaptation. The IO cares less about adaptation to a certain, and is therefore more willing to adapt to a country with a more extreme type, as long as decisions are coordinated. This increases the sensitivity of these decisions to a country's signal and thus, incentives to misrepresent are stronger because lying has a bigger

Figure 2: Coordination and the Value of Delegation ($s = 1$)



Note: The solid curve is welfare without delegation, while the dashed curve is welfare with delegation. The value of coordination β ranges from 0 *to* 1.

influence on policies. In turn, this increases the willingness to burn money. Additionally, with delegation, country 1 must send signals to influence d_1, while it could freely choose d_1 without delegation. This increases money burning incentives as well. Proposition 2 establishes equilibrium money burning strategies when authority is delegated.

Proposition 2. *In the equilibrium of the delegation game, country i burns*

$$b_i^D = \frac{2(1-\beta)\beta}{1+3\beta} f_i(\theta_i).$$

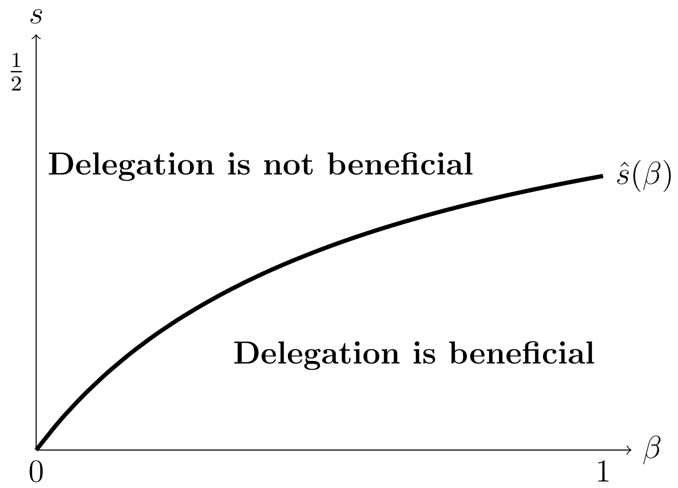
The Value of Delegation

Having characterized equilibrium strategies, we now compare the benefits of countries with and without an IO. This is determined by two parts: which decisions are taken and signaling costs. Proposition 3 disentangles each country's political and informational welfare, and provides conditions under which delegation is beneficial for a country.

Proposition 3. *Delegation increases political welfare for each country but decreases informational welfare. Delegation is beneficial for a country if and only if the level of uncertainty is sufficiently low relative to the value of coordination.*

In delegating authority, there is a trade-off between coordination and information trans-

Figure 3: Coordination, Uncertainty, and the Value of Delegation



Note: The curve illustrates the function $\hat{s}(\beta)$ for which delegation and no delegation are equally good from countries' perspectives ex ante.

mission. On the one hand, countries benefit from the IO as it induces additional coordination. On the other, countries need to transmit their information and we show that this is more costly if an IO is involved.

Figure 2 illustrates how high each country's political and informational welfare is under both institutions for varying values of β, assuming that $s = 1$. As the figure illustrates, both institutions are almost equivalent if β is close to the extremes, but overall, political welfare is higher under delegation, while informational welfare is lower.

The difference in performance between both institutions depends on how countries value coordination. If β is low, then there is little value in achieving coordinated policies, and neither country has an interest in influencing the other country. If β is high, preferences are highly aligned and both countries do not mind to give in to the other country, provided that both decisions are highly coordinated. If β is intermediate, however, then delegation leads to a significant increase in political welfare because IOs increase coordination, and it is valued by both countries. Additionally, the increase in money burning is also significant for intermediate β, as both countries have a greater stake in sending signals to obtain coordinated policies that are closer to each country's bliss point.

Proposition 3 formally establishes that delegation is beneficial if and only if the level of uncertainty s is sufficiently low relative to the importance of coordination. This means that, for a fixed value of β, there exists a value $\hat{s}(\beta)$ such that if $s < \hat{s}(\beta)$, countries prefer delegation, while if $s > \hat{s}(\beta)$, countries are better off without an IO. Figure 3 illustrates this graphically. There are two main takeaways. First, higher β makes countries gain more from delegation for a fixed level of uncertainty s, and for a fixed β, more uncertainty makes countries gain less from IOs. Thus, if countries value coordination to a large extent, they benefit from delegation for a wider range of values s. For countries to gain from delegation, it is necessary that they sufficiently value coordination but that there is not too much uncertainty. Second, if there is too much uncertainty, then countries never find it beneficial to delegate, even if coordination is highly important. This highlights the detrimental negative effect of delegation on signaling, especially when countries face a very uncertain environment.

The fact that IOs bring about more coordination and lead to more costly signaling is confirmed in Appendix section C.8. There, we alter the IO's preferences to give it any arbitrary weight on the coordination motive without caring about adaptation. This leads to both more coordination and more money burning by each country. That is, countries have stronger incentives to influence the IO when decisions are more coordinated.

Countries may take decisions both when IOs are and are not involved. As several examples illustrate, IOs can be quite successful in improving coordination. For instance, in the 1960s, European countries failed to develop regional color TV standards, while in contrast, they were much more successful in the European Community in the 1980s to set HDTV standards (Austin and Milner, 2001). Our results confirm that IOs are beneficial for international coordination. The examples also showed that, due to international interdependence, countries have incentives to influence each other. This influence may take several forms, by either sending a letter as in the green energy example, or by engaging in costly activities such as organizing conferences or subsidizing firms. It is more difficult to empirically establish what effect IOs have on the transmission of information. Our results provide some indication that countries incur greater costs with influence activities when IOs are involved in decision-making. For instance, the French may give more subsidies and support for nuclear energy than it would otherwise have done, partially to influence other EU member states. That is, France has an individual incentive to invest in nuclear energy, but may increase it even more, given that other EU member states are watching. Our model predicts that, absent IOs, decisions would be less coordinated, but France would have fewer incentives to influence other countries. This means that, although countries may gain from additional coordination through IOs, they pay an extra price in influencing other countries.

Scope and Applications

With our main results in hand, and prior to presenting the extensions section, we briefly discuss several policy areas which our model speaks to. Our discussion serves to illustrate how countries may signal their information in practice beyond simple communication, and how countries trade off coordination and adaptation in the presence and absence of IOs. The examples help illustrate different elements of the model and results.

Television Standards

Consider the selection of television standards by multiple countries. Countries benefit from coordinating and imposing similar technological standards due to network externalities in consumption and scale economies in production. Standards reduce uncertainty, as they provide firms with the security that technologies are unlikely to be abandoned in the future due to a subsequent development of incompatible standards. As a result, international coordination leads to higher levels of investments, which directly benefits firms, but indirectly governments as well. While countries have a common benefit in attaining convergent standards, they also want to promote their national audiovisual and electronic industries, which pushes preferences towards one standard over another. As a result, countries have incentives to persuade other countries to set standards that benefit their domestic firms.

An example of this is the adoption of color television standards in Europe. France exerted significant effort to influence other countries by, e.g., assistance in marketing surveys and research, assistance in sales, subsidies to firms, promotional tours, exhibitions, fairs, and demonstrations. (Crane, 1978).[18] France may partly use these signals to convince other countries to use the French preferred standard. All such methods go further than communication and the more costly these actions are, the more strongly they signal the intensity of French preferences, and the more influential they could be. Although firms are directly gaining from coordinated standards, governments indirectly gain as well, and therefore have aligned incentives with firms to try to influence other countries.

Technological Standards

More recently, technological standards in the Fourth Industrial Revolution[19] have been extensively debated. Governments help in the diffusion and adoption of technologies developed by domestic companies to increase demand. Countries, however, have different industries, which leads to diverging preferences on optimal technologies. Examples include electric cars, internet of things (IoT) platforms, autonomous vehicles, bike-sharing, payment systems, facial recognition, artificial intelligence, data protection law, and so on. Efforts to develop their technologies and influence countries to adopt those technologies include Germany spending 200 million euros with the program 'Industrie 4.0' and China spending almost 200 billion euros with the program 'Made in China 2025'.[20]

In addition, countries attempt to influence IOs in the selection of standards. In Europe, Germany pushes relatively more than other countries to influence the common standard in technologies set within the EU. China behaves similarly, influencing standards selected by the ITU (International Telecommunication Union)[21] and the ISO (International Organisation for Standardisation). Countries such as the United States and China have also tried to influence other countries' standards directly without IOs, especially in developing

[18] Angulo, Calzada and Estruch (2011) discusses similar behavior in Latin American standards for digital television, where countries wanted to influence their neighbors to adopt similar standards.

[19] This refers to the automation of manufacturing and industrial practices, using modern smart technologies. See https://www.foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution (accessed 11/22/2020).

[20] See https://jacobinmag.com/2020/11/digitalization-european-union-market-us-china-tech (accessed 11/22/2020).

[21] Standards ratified in the ITU are commonly adopted as policy by developing nations in Africa, the Middle East and Asia. See https://www.ft.com/content/c3555a3c-0d3e-11ea-b2d6-9bf4d1957a67 (accessed 11/22/2020).

countries, as they lack the resources to develop standards themselves.[22]

The GSM Standard

Another example that our model may speak to is the GSM standard in Europe. Pelkmans (2001) mentions how in the early 1980s, Western Europe was highly uncoordinated with respect to mobile communication. Each country had their own monopolistic market with different standards. This was costly for monopolists—and thus also for the governments due to decreased tax revenue—because they could not benefit from the wider European market. For the production of all types of equipment, scale economies were necessary to do so profitably; as business demand was simply too small in each individual country. As a result, understanding the benefits of coordination, countries were able to successfully let European institutions be involved in setting the common GSM standard. Starting in 1985, the European Commission took an active role in promoting the GSM standard, and was successful over the following years. This example illustrates how IOs may help countries to adopt similar standards as well, again to the direct benefit of national firms and governments.

Extensions

In this section, we study several extensions in which we relax the main assumptions of the model. Some of the extensions analyze asymmetric environments, while others are concerned with different decision-making structures.

Bargaining in International Organizations

In the baseline delegation game, we assumed that a separate player—the IO—weighed both countries' utilities equally and took decisions on their behalf. We now analyze a variation of the model, the *international bargaining* game, in which countries bargain over policy. This is to investigate whether the results are purely driven by a focus on centralized IOs rather than member-led IOs. In the latter, it is more natural to think of

[22]https://www.ft.com/content/0c91b884-92bb-11e9-aea1-2b1d33ac3271 (accessed 11/22/2020).

countries still being the relevant actors in making decisions. Still, IOs perform a particular task in helping with cooperation. The idea is that, once countries are members of an IO, it becomes easier to commit to agreed upon policies. That is, once a proposal is accepted, both countries are bound to stick to the agreement. After Nature has drawn countries' types θ_1 and θ_2 and countries burned money, the timing is now as follows:

1. Nature selects country 1 as the proposer with probability $p \in [0,1]$ and country 2 with probability $1-p$.
2. The proposer observes signals and offers (d_1, d_2, T), where $T \in \mathbb{R}$ is a transfer.
3. The non-proposer observes (d_1, d_2, T) and the signals, and accepts or rejects.
4. If the offer is accepted, the outcome is (d_1, d_2, T), otherwise, countries' decisions are as if there were no delegation. Payoffs are equivalent to the main model except for the transfer T, which is valued linearly.

Our equilibrium concept is the same as in the main model, with the added restriction that we focus on fully separating equilibria through burned money. Solving backward, if the offer is rejected, countries make decisions as in the no-delegation game, established in Lemma 2. Hence, in contemplating whether to accept or reject an offer, a country compares the offer (d_1, d_2, T) with its outside option of the no-delegation game. Knowing this, the proposer wishes to maximize his utility subject to the other country's acceptance. A first result is that the proposer always offers (d_i^D, d_j^D) as the IO does in the delegation game, as in Lemma 3. The proposer then sets the transfer T so that the other country is made indifferent between the offered (d_i^D, d_j^D) and its outside option from the no-delegation game. Thus, it is beneficial to be the proposer. Equilibrium decisions (d_i^D, d_j^D) are the same regardless of the proposing country, but the proposer is able to extract a transfer from the other country. The following proposition establishes the results.

Proposition 4. *In the international bargaining game, the following statements are true:*

1. *If countries have equal bargaining power, equilibrium strategies are equivalent to the delegation game.*
2. *For every distribution of bargaining power, countries burn more than in the no-*

delegation game.

3. *If a country obtains more bargaining power, it burns more money.*

Thus, if $p = \frac{1}{2}$, both countries have equal probability of being the proposer and we show that decision-making strategies and money burning strategies are equivalent to those in the delegation game. This provides a formal justification for the setup of the delegation game. In addition, for all $p \in [0, 1]$, international bargaining leads to more aggregate money burning than in the no-delegation game. This highlights how even with asymmetric bargaining power, countries have stronger incentives to burn money. Finally, we also show that a country burns more money if its bargaining power increases. For example, if country 1 is the proposer, then if it signals that it has an extreme type, it reveals negative information about the other country's outside option. This allows country 1 to extract a greater transfer from country 2 than otherwise, which determines money burning strategies. Alternatively, if country 2 is the proposer, then country 1 wants to mimic extreme types to induce the proposer to offer (d_1, d_2) that are closer to 1's bliss point. These different incentives to burn money generate the result that lying is more attractive if a country is the proposer than if it is not.

The Importance of Enforcement

Countries generally remain autonomous over policy-making even if IOs are involved. Hence, it is an open question how the enforcement capabilities of IOs affect strategic behavior of countries. We extend the delegation game by allowing countries to deviate from the IO's (recommended) decisions. That is, at some cost $c \geq 0$, which measures the IO's capabilities to enforce policies, each country may deviate from what the IO decides. The game is now as follows after Nature draws types and countries burned money:

1. The IO sets policies (d_1, d_2).

2. Countries 1 and 2 individually make decisions. If country $i = 1, 2$ deviates from d_i, it pays cost $c \geq 0$.

In observing a proposal, countries contemplate whether they want to deviate from it.

Unless the IO chooses decisions as in the no-delegation game (which is the unique equilibrium without an IO), there is always a profitable deviation if there is no cost of deviating. Hence, if the IO anticipates potential deviations with low c, it wants to prevent this. Alternatively, if c is high enough, it can simply take decisions as in the delegation game in which no deviations were possible. The lower is c, the more the IO must accommodate countries to prevent them from deviating.

In turn, the value of c also has an effect on money burning strategies as countries anticipate less coordinated decisions from the IO if c is low. Put differently, if c is low, we obtain decisions and money burning strategies closer to the no-delegation game. If c is high, we obtain that it is equivalent to the delegation game with higher levels of money burning. The following proposition summarizes the results of this extension.

Proposition 5. *In every equilibrium in which the IO makes a proposal that is accepted and from which neither country will deviate, the IO makes proposals as follows: If* $c \leq \frac{(1-\beta)^2\beta^2\mathbb{E}_{IO}[\theta_1-\theta_2]^2}{(1+3\beta)^2}$*, then* $d = \left(d_1^{ND} + \frac{\sqrt{c}}{1+\beta}, d_2^{ND} - \frac{\sqrt{c}}{1+\beta}\right)$*. Otherwise, decisions are as in the delegation game. Further, countries' money burning strategies are weakly increasing in the level of enforcement* c*.*

As can be seen from the result above, low values of c leads to equilibrium structures similar to the no delegation game, whereas high values of c generates equilibrium strategies that are similar to the delegation game. We observe that our previous results are two extreme cases in terms of enforcement capabilities. Another salient aspect is that the level of enforcement necessary to generate decisions as in the delegation model is greater for higher levels of disagreement. That is, when θ_1 and θ_2 are very different, countries have different bliss points, and the IO must have enough enforcement capabilities c to sustain coordinated decisions.

Preference Alignment and Delegation

We now evaluate how preference misalignment affects gains from delegation. Having established the equivalence between the international bargaining game and the delegation game, we return to the latter. We model varying levels of disagreement by altering $\mathbb{E}^0[\theta_1]$

and $\mathbb{E}^0[\theta_2]$, the means of the type distributions, but keeping them symmetric around 0. Formally, $\mathbb{E}^0[\theta_2] = -\mathbb{E}^0[\theta_1] \geq 0$, and both countries are in expectation equally aligned to the IO. The difference between the two means is then the level of disagreement, defined as $\Delta := \mathbb{E}^0[\theta_2] - \mathbb{E}^0[\theta_1]$. In the baseline model, $\Delta = 2$, but now we allow for any $\Delta \geq 0$ and evaluate how delegation incentives are shaped by Δ. We assume that the level of uncertainty is bounded above with $s \in [0, \Delta)$.

We find that alignment decreases the gains from delegation. The reason is that countries already coordinate their decisions to a large extent when there is little disagreement between them. Any marginal increase in coordination is less likely to outweigh the increased costs of money burning when countries have more similar preferences. As a result, there is a larger scope for delegation to be beneficial for higher values of disagreement. IOs have a larger marginal contribution in achieving more coordination when there is more disagreement, and due to the quadratic nature of payoffs, this additional coordination is relatively valuable. Proposition 6 establishes the result.

Proposition 6. *The relative benefit of delegation vis-à-vis no delegation increases in* Δ.

This result illustrates that countries with more dissimilar standards have more to gain from the involvement of IOs. For instance, EU member states with similar technological standards and preferences would already be able to update and adapt their standards to accommodate multiple technologies without formal institutions, while countries such as Japan and Germany with less similar standards would have more to gain, even though delegation is unlikely for other reasons. If countries find coordination valuable but have different domestic circumstances, then their most preferred policies differ, and to achieve the gains from coordination, formal institutions have greater influence on bringing policies and standards closer together. Note, however, that as the previous extension has shown, enforcement becomes a bigger hurdle for countries to cross.

Asymmetric Uncertainty

We now turn to an extension to study equilibrium behavior with asymmetric uncertainty. More formally, there is one-sided incomplete information: country 2 has private information, while country 1's type θ_1 is observable to both countries. An interpretation of this extension could be that it is already clear what one country wants, but not yet for the other. For instance, if a country's economy has for a long history relied on certain technologies in producing energy, then other countries know what its preferences are. For developing countries or countries with volatile political environments, other countries may start off with some uncertainty. Alternatively, in developing standards, the country with a public type may be already have some first mover advantage and locked in certain standards, while the one with a private type is a second mover.

In the model, the country without private information has no incentive to burn money because it does not affect decisions. On the flip side, the country with private information still has an incentive to burn money, but can now condition its strategy on what it has observed about the other country. Its incentives to incur signaling costs are stronger when the marginal benefit of misrepresenting is higher. Whenever the other country's type suggests it will make a decision that is far from one's own bliss point, then there is more reason to signal that one's own type is higher. The reason is that a marginal change toward more coordinated policies is more valuable if preferences are less congruent.

In evaluating whether delegation is jointly beneficial, we find that there is a wider range of parameter values for which delegation is optimal. This effect is determined by the fact that now only one country burns money, which implies that the benefits of coordination are more likely to outweigh the costs of deteriorated signaling. Proposition 7 shows this result formally.

Proposition 7. *Delegation is jointly beneficial if and only if the level of uncertainty is sufficiently low such that $s < \check{s}(\beta)$, where $\check{s}(\beta) > \hat{s}(\beta)$.*

In this extension, country 1 always benefits from delegation because there are gains from increased coordination but no losses due to costly signaling. As a result, the only restriction is on country 2, who may lose from delegation. There is thus a wider range of parameters for which delegation may be jointly beneficial, as the losses that country 2 incurs may be offset by the gains made by country 1.

Selection of IO's Preferences

We now study how the gains from delegation depend on countries' potentially heterogeneous values of coordination β. This extension serves to capture a situation with a large and a small country. In a relative sense, the larger country such as China or the US cares more about adjusting decisions to domestic conditions, while a smaller country's economy is more strongly affected by decisions made in the large country and thus faces a larger cost if decisions are uncoordinated.

We study the extreme case and assume country 1 does not value coordination and has policy preferences of $\pi_1(d_1, \theta_1) = -(d_1 - \theta_1)^2$.[23] Country 2, however, still values coordination with weight $\beta_2 \in (0, 1)$ and has preferences as in the main model. We study preferences of the IO that are still a weighted average of both countries' interests i.e.,

$$u_{IO}(\theta_1, \theta_2) = \alpha \left[-(d_1 - \theta_1)^2\right] + (1 - \alpha) \left[-(1 - \beta_2)(d_2 - \theta_2)^2 - \beta_2(d_1 - d_2)^2\right].$$

In the benchmark we assume $\alpha = \frac{1}{2}$ and provide the following result.

Lemma 4. *Country 1 never prefers to delegate while country 2 always prefers to delegate. There exists an inverted u-shaped function $\tilde{s}(\beta)$ such that, if $s \leq \tilde{s}$, then delegation generates joint benefits.*

Thus, although country 1 would lose from delegation, country 2 gains more and could compensate for this loss by sending transfers as long as the level of uncertainty is sufficiently low. The reason for the non-monotonic effect becomes apparent by contrasting two extreme situations. If country 2 cares very little about coordination, with $\beta_2 \approx 0$, then the positive effects of the increase in coordination due to delegation is unlikely to outweigh the costs of money burning, even with little uncertainty. In the other extreme where country 2 finds coordination highly important, with $\beta_2 \approx 1$, country 2 is already willing to coordinate to a large extent with the other country, again implying that delega-

[23] We study the case where country 1 does not value coordination ($\beta_1 = 0$) to understand the general logic of the results. The analysis follows through when both countries value coordination positively with $\beta_1 > 0$ and $\beta_2 > 0$.

tion is barely beneficial. Increased coordination is only valuable for intermediate values of β_2, and may lead to beneficial delegation for a wider range of uncertainty.

Another way to compensate country 1 is to alter the allocation of authority in the IO. We now study how the joint benefits from delegation can be maximized by selecting $\alpha \in [0, 1]$, which is country 1's weight in the IO. An increase in α grants country 1 more authority, shifting decisions in 1's favor, and also affects the signals that countries send. Proposition 8 establishes our result in this section. The results depend crucially on the amount of uncertainty, s, and the importance that the small country places on coordination, β_2. If the goal of the IO is to generate the largest amount of joint benefits, then the share of authority by country 1 is increasing in the level of uncertainty.

Proposition 8. *The institution that maximizes joint benefits always gives weakly more authority to country* 1 *with* $\alpha \geq \frac{1}{2}$. *Further, there exists a function* $s^*(\beta_2) < 1$ *such that if there is more uncertainty than* $s^*(\beta_2)$, *then all authority is in the hands of country* 1 ($\alpha = 1$). *If there is less uncertainty than* $s^*(\beta_2)$, *then* α *is increasing in the amount of uncertainty.*

There are two factors that affect the total gains from delegation. First, each country's payoffs that are determined by equilibrium decisions. Given that institutions that maximize the total gains for both countries weigh the welfare of them both equally, it implies that if there is no uncertainty, countries should have equal authority. This guarantees that decisions are taken that weigh both countries' interests equally. The second factor is informational welfare. Country 1 never burns money because it has no interest in changing country 2's behavior. As a result, the only costly signals are sent by country 2. With more uncertainty, this part affects the gains and losses from delegation the most, and by giving country 1 more authority, country 2 knows that its signals have less influence on decisions, reducing incentives to send costly signals. With too much uncertainty, it is optimal to give all authority to country 1.

Limited Discretion

In this extension we ask how limiting the IO's discretion affects countries' gains from delegation? Formally, the IO is restricted to choose the same decision for both countries $d = d_1 = d_2$ and has to select d from a delegation interval. We assume that this interval is symmetric about 0 with $[-\ell/2, \ell/2]$, which has length $\ell \geq 0$. Without discretion ($\ell = 0$), the IO is forced to select $d = 0$, while if $\ell = \infty$, the IO has unlimited discretion.

The IO's decision-making strategy is the following function of its beliefs about θ_1 and θ_2. The IO's ideal policy based on its beliefs is equal to $\hat{d}_{IO} := \frac{1}{2}\left[\mathbb{E}_{IO}(\theta_1|b_1, m_1) + \mathbb{E}_{IO}(\theta_2|b_2, m_2)\right]$. If this ideal policy falls within the IO's delegation interval, then it is the outcome, otherwise the policy is its lower ($-\ell/2$) or upper bound ($\ell/2$).

Changing the IO's discretion not only alters decisions but also countries' signals. When the IO has relatively more discretion, signals have a greater influence on decisions, which increases incentives to burn money. In selecting the IO's level of discretion, there is a trade-off between getting decisions that are more tailored to countries' domestic circumstances, and incurring costs to transmit information through money burning.

We evaluate how the length of the IO's delegation interval affects each country's benefits of delegation. With little uncertainty, it is best to give the IO no discretion. When the level of uncertainty increases, the size of the delegation interval grows. The interval's length ℓ depends negatively on the amount of disagreement between the two countries. Finally, the length of the delegation interval does not depend on the value of coordination β. This is because the IO must take fully coordinated decisions, implying that there is only a trade-off between the benefits of adaptation and the cost of burned money.

Proposition 9. *In each country's most preferred institution, the length of the delegation interval ℓ increases in the level of uncertainty s and decreases in the amount of* *disagreement* Δ, *where*

$$\ell(s,\Delta) = \begin{cases} 0 & \textit{if } 0 \leq s \leq \sqrt{3}(\Delta/2), \\ \frac{s}{3} - \frac{\Delta^2}{4s} & \textit{if } \sqrt{3}(\Delta/2) < s \leq \Delta. \end{cases}$$

The results further emphasize our earlier findings about how IOs negatively impact signaling. With even more coordination after delegation, incentives to burn money are even stronger than in the baseline model. This makes it necessary to limit the IO's discretion to dampen money burning incentives. Further, we show that each country's most preferred length of the delegation interval increases in s because it increases the potential for both countries to have the same type $\theta_1 = \theta_2$. When type spaces do not overlap ($s < \Delta/2$), it is never optimal to give the IO any discretion. Also, as shown earlier, greater disagreement makes money burning incentives stronger because there is more to gain from influencing the IO's decision, further increasing the benefits of limited discretion.

This extension can be applied to the selection of standards as well. Countries may cooperate in, e.g., the ITU to facilitate common standards for HDTV. Japan attempted to influence the world standard against the ones used in Europe, but the Europeans coordinated within their own institutions. Our results demonstrate that, as the ITU recommends a single standard, countries have much stronger incentives to try and influence the standard. Then, the more limited the ITU's discretion is in influencing standards, the less incentives countries have to send costly signals in transmitting information. Beyond the ITU, the results help in rethinking the design of other, future IOs as well, where limiting discretion can be helpful to curb countries' influence activities. Although this may prevent IOs from adapting policies to countries' circumstances, this loss of limited discretion may be outweighed by the gain in weakening costly signaling.

Discussion and Conclusion

We have studied a model of international coordination in which countries have information that IOs lack. IOs help countries coordinate more but they also induce informational losses due to increased incentives to send costly signals. The main result is that delegation is beneficial for countries as long as they sufficiently value coordination relative to the amount of uncertainty about other countries' preferences.

In contrast to established theories that emphasize the benefits of IOs for sharing and pooling information (Koremenos, 2008), our novel result is that IOs perform worse in an informational sense. Countries are equally able to transmit information with and without IOs in our setup of coordinated adaptation. This contradicts previous theories that unequivocally emphasize the benefits of IOs in generating more information for its member states than they would have obtained otherwise. Moreover, signaling is costlier when IOs are involved. Thus, although countries may prefer that IOs are involved in decision-making because international policies are more coordinated, the road to get there has significant costs. In our model, these costs could take many forms, such as subsidies, research and development, costly investments, and lobbying. The main component of these activities, however, is that they are costly and help persuade other countries.

We predict that similar patterns hold true in a model without money burning, but in which countries may incur costs in other ways. For example, if two countries are incompletely informed and engage in bargaining over time, then an established result is that countries may delay decision-making to signal their resolve or preference intensity (Rubinstein, 1985; Abreu and Gul, 2000). Thus, instead of actively engaging in costly activities, delaying decision-making may act a similar tool to send other countries a signal about preferences, which helps if cheap talk is only partially informative. In a similar vein, Urpelainen (2012) studies a model where international bureaucrats are influenced by different countries to get preferred policies. Although this is a model with complete information, it shows that countries engage in costly activities to influence decisions. If, as in our model, countries would be incompletely informed about other countries, this would exacerbate the problem because costly activities are not merely directly influential, they also have an indirect signaling value. Countries that have more to gain from influencing policies are more willing to engage in costly actions.

The results have several implications for countries in the context of international coordination. First, we should observe more successful coordination when IOs are involved in decision-making. IOs are able to push countries to coordinate, for example, in selecting common product standards. At the same time, however, we predict that countries exert more effort in trying to get policies in their favor if IOs are involved. In this sense, countries engage in more wasteful activities to ensure that preferred decisions are taken.

In two main robustness checks, we show that information sharing and decision-making occurs in a qualitatively similar way if countries bargain over policies. That is, if countries have some probability of drafting a proposal that must be accepted by the other country, this can generate the same results compared to when a separate entity makes decisions on their behalf. Second, a driving force of an IO's impact on information sharing and decision-making is its capability to enforce decisions. Going from weak to strong IOs, policy outcomes become more coordinated, and hence, countries need to send costlier signals to ensure that all information is transmitted. The enforceability of decisions is especially salient when countries disagree more about which policies work best for their respective economies.

We provide several other implications for institutional design beyond the simple binary delegation comparison. First, if countries are less aligned, they actually gain more from the involvement of IOs. Thus, we may observe coordination without formal institutions if countries are aligned, as IOs only generate benefits when countries need institutions to enforce coordination. Second, when there is relatively less uncertainty about one country's preferences than the other, this may generate a wider scope for IOs to be beneficial for both countries. Third, larger countries that care relatively less about coordination must be compensated for them to gain from IOs. That is, either smaller countries must compensate through monetary benefits (similar to Urpelainen (2012)) or allow larger countries to have more authority in IOs. Fourth, from the perspective of countries, IOs may be best designed with limited discretion. Although this prevents policies from being more tailored to countries' circumstances, the benefits of reduced costly signaling is more likely to outweigh this, especially when countries are less aligned.

Naturally, as the literature shows, there are several potential reasons for why IOs are not organized to maximize the joint benefits of member states. Instead, countries bargain over institutional design, and those that have better outside options often have a greater say in the design process (Johns, 2007; Lipscy, 2017). Additionally, countries are not in complete control over institutional design in IOs, and international bureaucrats play an important role (Johnson, 2013, 2014; Johnson and Urpelainen, 2014). Thus, even if limited discretion is best from countries' perspectives, we may not observe it in practice if bureaucrats have bargaining power.

We study the effect of formal institutions on informational issues in a framework where countries trade off coordination and adaptation. Future work could help analyze similar issues in different frameworks. For example, if countries face a free-rider problem (Kenkel, 2019), it is an open question if IOs are still beneficial when countries must transmit information prior to policy-making. In different policy areas, preferences may be determined by a competition motive rather than a coordination motive (Lazer, 2001). Our results, however, indicate that informational issues in international policy-making cannot be underestimated, especially in highly uncertain environments, and that IOs may make them worse rather than solve them.

References

Abbott, Kenneth W and Duncan Snidal. 1998. "Why States Act Through Formal International Organizations." *Journal of Conflict Resolution* 42(1):3–32.

Abreu, Dilip and Faruk Gul. 2000. "Bargaining and Reputation." *Econometrica* 68(1):85–117.

Alonso, Ricardo, Wouter Dessein and Niko Matouschek. 2008. "When Does Coordination Require Centralization?" *American Economic Review* 98(1):145–79.

Angulo, Jorge, Joan Calzada and Alejandro Estruch. 2011. "Selection of Standards for Digital Television: The battle for Latin America." *Telecommunications Policy* 35(8):773–787.

Austen-Smith, David and Jeffrey S. Banks. 2000. "Cheap Talk and Burned Money." *Journal of Economic Theory* 91(1):1–16.

Austin, Marc T and Helen V Milner. 2001. "Strategies of European Standardization." *Journal of European Public Policy* 8(3):411–431.

Barnett, Michael N and Martha Finnemore. 1999. "The Politics, Power, and Pathologies of International Organizations." *International Organization* 53(4):699–732.

Bendor, Jonathan, Amihai Glazer and Thomas Hammond. 2001. "Theories of Delegation." *Annual Review of Political Science* 4(1):235–269.

Bernheim, B Douglas and Sergei Severinov. 2003. "Bequests as Signals: An explanation for the equal division puzzle." *Journal of Political Economy* 111(4):733–764.

Bradley, Curtis A and Judith G Kelley. 2008. "The Concept of International Delegation." *Law and Contemporary Problems* 71(1):1–36.

Carnegie, Allison and Austin Carson. 2019. "The Disclosure Dilemma: nuclear intelligence and international organizations." *American Journal of Political Science* 63(2):269–285.

Carson, Austin and Alexander Thompson. 2014. "The Power in Opacity: Rethinking Information in International Organizations.".

Chapman, Terrence L. 2007. "International Security Institutions, Domestic Politics, and Institutional Legitimacy." *Journal of Conflict Resolution* 51(1):134–166.

Chen, Ying, Navin Kartik and Joel Sobel. 2008. "Selecting Cheap-Talk Equilibria." *Econometrica* 76(1):117–136.

Cho, In-Koo and David M Kreps. 1987. "Signaling Games and Stable Equilibria." *The Quarterly Journal of Economics* 102(2):179–221.

Crane, Rhonda J. 1978. "Communication Standards and the Politics of Protectionism: The case of colour television systems." *Telecommunications Policy* 2(4):267–281.

Crawford, Vincent P. and Joel Sobel. 1982. "Strategic Information Transmission." *Econometrica* 50(6):1431–1451.

Cressman, Ross and Maria Gallego. 2009. "On the Ranking of Bilateral Bargaining Opponents." *Mathematical Social Sciences* 58(1):64–83.

Crombez, Christophe, Martijn Huysmans and Wim Van Gestel. 2017. "Choosing an Informative Agenda Setter: The appointment of the Commission in the European Union." *European Union Politics* 18(2):145–167.

Fang, Songying. 2008. "The Informational Role of International Institutions and Domestic Politics." *American Journal of Political Science* 52(2):304–321.

Fang, Songying and Randall W Stone. 2012. "International Organizations as Policy Advisors." *International Organization* 66(04):537–569.

Fearon, James D. 1995. "Rationalist Explanations for War." *International organization* 49(03):379–414.

Hawkins, Darren G, David A Lake, Daniel L Nielson and Michael J Tierney. 2006. *Delegation and Agency in International Organizations*. Cambridge University Press.

Hooghe, Liesbet and Gary Marks. 2015. "Delegation and Pooling in International Organizations." *The Review of International Organizations* 10(3):305–328.

Johns, Leslie. 2007. "A Servant of Two Masters: Communication and the selection of international bureaucrats." *International Organization* 61(02):245–275.

Johnson, Tana. 2013. "Looking beyond States: Openings for international bureaucrats to enter the institutional design process." *The Review of International Organizations* 8(4):499–519.

Johnson, Tana. 2014. *Organizational Progeny: Why governments are losing control over the proliferating structures of global governance*. Transformations in Governance.

Johnson, Tana and Johannes Urpelainen. 2014. "International Bureaucrats and the Formation of Intergovernmental Organizations: Institutional design discretion sweetens the pot." *International Organization* 68(1):177–209.

Karamychev, Vladimir and Bauke Visser. 2016. "Optimal Signaling with Cheap Talk and Money Burning." *International Journal of Game Theory* 46(3):1–38.

Kartik, Navin. 2009. "Strategic Communication with Lying Costs." *The Review of Economic Studies* 76(4):1359–1395.

Kenkel, Brenton. 2019. "The Efficacy of Cheap Talk in Collective Action Problems." *Journal of Theoretical Politics* 31(3):370–402.

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

Keohane, Robert O. 1984. *After Hegemony: Cooperation and discord in the world political economy*. Princeton University Press.

Kolotilin, Anton and Hongyi Li. 2021. "Relational Communication." *Theoretical Economics* 16(4):1391–1430.

Koremenos, Barbara. 2008. "When, What, and Why do States Choose to Delegate?" *Law and Contemporary Problems* 71(1):151–192.

Koremenos, Barbara, Charles Lipson and Duncan Snidal. 2001. "The Rational Design of International Institutions." *International Organization* 55(04):761–799.

Lazer, David. 2001. "Regulatory Interdependence and International Governance." *Journal of European Public Policy* 8(3):474–492.

Lipscy, Phillip Y. 2017. *Renegotiating the World Order: Institutional Change in International Relations*. Cambridge University Press.

McAllister, Jordan H and Keith E Schnakenberg. 2021. "Designing the Optimal International Climate Agreement with Variability in Commitments." *International Organization* pp. 1–18.

Morrow, James D. 1994. "Modeling the Forms of International Cooperation: Distribution versus information." *International Organization* 48(3):387–423.

Pelkmans, Jacques. 2001. "The GSM Standard: explaining a success story." *Journal of European Public Policy* 8(3):432–453.

Rantakari, Heikki. 2008. "Governing Adaptation." *The Review of Economic Studies* 75(4):1257–1285.

Rubinstein, Ariel. 1985. "A Bargaining Model with Incomplete Information about Time Preferences." *Econometrica* pp. 1151–1172.

Stone, Randall W. 2009. Institutions, Power, and Interdependence. In *Power, interdependence, and nonstate actors in world politics*. Princeton University Press pp. 31–49.

Thompson, Alexander. 2015. *Channels of Power: The UN Security Council and US statecraft in Iraq*. Cornell University Press.

Toulemonde, Eric. 2013. "A Welfare Analysis of the Principle of Mutual Recognition." *European Economic Review* 60:1–16.

Urpelainen, Johannes. 2012. "Unilateral Influence on International Bureaucrats: An international delegation problem." *Journal of Conflict Resolution* 56(4):704–735.

Voeten, Erik. 2019. "Making Sense of the Design of International Institutions." *Annual Review of Political Science* 22:147–163.

Supplementary Information for "International Coordination and the Informational Rationale for Delegation"

Online Appendix

Table of Contents

A Model Setup and Equilibrium

A.1 Equilibrium Concept

In the no-delegation game, country i's strategy is (i) a mapping from types to signals $(b_i^{ND}, m_i^{ND}) : \Theta_i \to \mathbb{R}_+ \times \mathbb{R}_+$ and (ii) a mapping from country i's type, country i's and j's signals to decisions $d_i^{ND} : \Theta_i \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$.
In the delegation game, country i's strategy is a mapping from types to signals $(b_i^D, m_i^D) : \Theta_i \to \mathbb{R}_+ \times \mathbb{R}_+$; and for the IO it is a mapping from the signals of both countries to decisions (d_1^D, d_2^D) with $d_i^D : \mathbb{R}_+^2 \times \mathbb{R}_+^2 \to \mathbb{R}$. Let $\mu_i^{ND}(b_j, m_j) \in \Delta(\Theta_j)$ be country i's posterior belief about country j's type after observing (b_j, m_j) and $\mu_{IO}^D(b_i, m_i, b_j, m_j) \in \Delta(\Theta_i) \times \Delta(\Theta_j)$ the IO's posterior beliefs about countries i and j's types after observing (b_i, m_i, b_j, m_j). Denote $b^{\mathcal{I}} = (b_1^{\mathcal{I}}, b_2^{\mathcal{I}})$, $m^{\mathcal{I}} = (m_1^{\mathcal{I}}, m_2^{\mathcal{I}})$, $d^{\mathcal{I}} = (d_1^{\mathcal{I}}, d_2^{\mathcal{I}})$, $\mu^{ND} = (\mu_1^{ND}, \mu_2^{ND})$ and $\mu^D = \mu_{IO}^D$.

Formally, a *perfect Bayesian equilibrium* (PBE), and from now on an *equilibrium*, is a tuple $(b^{\mathcal{I}}, m^{\mathcal{I}}, d^{\mathcal{I}}, \mu^{\mathcal{I}})$ where $(b^{\mathcal{I}}, m^{\mathcal{I}}, d^{\mathcal{I}})$ is sequentially rational given $\mu^{\mathcal{I}}$ and $\mu^{\mathcal{I}}$ is Bayesian consistent with $(b^{\mathcal{I}}, m^{\mathcal{I}})$.

In the no-delegation game, (b^{ND}, m^{ND}, d^{ND}) is *sequentially rational* given μ^{ND} if

For each θ_i,

$$\left(b_i^{ND}(\theta_i), m_i^{ND}(\theta_i)\right) \in \text{argmax}_{(b_i,m_i)} \mathbb{E}_i^0\Big[u_i\Big(d_i^{ND}\big(\theta_i, b_j^{ND}(\theta_j), m_j^{ND}(\theta_j)\big), d_j^{ND}\big(\theta_j, b_i, m_i\big), \theta_i, b_i\Big)\Big].$$

For each θ_i, b_j and m_j,

$$d_i^{ND}(\theta_i, b_j, m_j) \in \text{argmax}_{d_i} \mathbb{E}_i\Big[\pi_i\Big(d_i, d_j^{ND}\big(\theta_j, b_i^{ND}(\theta_i), m_i^{ND}(\theta_i)\big), \theta_i\Big)|b_j, m_j\Big].$$

In the delegation game, (b^D, m^D, d^D) is *sequentially rational* given μ^D if

For each θ_i,

$$(b_i^D(\theta_i), m_i^D(\theta_i)) \in \text{argmax}_{(b_i,m_i)} \mathbb{E}_i^0\Big[u_i\Big(d_i^D\big(b_i, m_i, b_j^D(\theta_j), m_j^D(\theta_j)\big), d_j^D\big(b_i, m_i, b_j^D(\theta_j), m_j^D(\theta_j)\big), \theta_i, b_i\Big)\Big].$$

For each b_i, m_i, b_j and m_j,

$$(d_i^D(b_i, m_i, b_j, m_j), d_j^D(b_i, m_i, b_j, m_j)) \in \text{argmax}_{(d_i,d_j)} \mathbb{E}_{IO}\Big[u_A\Big(d_i, d_j, \theta_i, \theta_j\Big)|b_i, m_i, b_j, m_j\Big].$$

In both cases, $\mu^{\mathcal{I}}$ is *Bayesian consistent* with $(b^{\mathcal{I}}, m^{\mathcal{I}})$ if $\mu_i^{\mathcal{I}}$ is the conditional probability distribution of θ_j given (b_j, m_j) derived from the joint distribution over $\Theta_i \times \mathbb{R}_+ \times \mathbb{R}_+$ that the prior distribution and $(b_j^{\mathcal{I}}, m_j^{\mathcal{I}}) : \theta_j \to \mathbb{R}_+ \times \mathbb{R}_+$ induce.

A.2 mD1-Refinement

We adopt the monotonic D1 refinement (Bernheim and Severinov (2003)). The D1 criterion (Cho and Kreps, 1987) says that after observing an out-of-equilibrium signal, players should not believe it is a type θ_i if there is some other type θ_i' who would strictly prefer to deviate for any response from the players that type θ_i would weakly prefer to deviate for. The monotonicity requirement is that higher types use weakly higher (or weakly lower) costly signals and posterior beliefs are monotonic as a function of these costly signals (including out-of-equilibrium signals).

The D1 refinement rules out the possibility of pooling intervals. The monotonicity requirement gives some order to posterior beliefs regarding signals that rule out pooling from isolated types. In sum, both conditions allow us to have an equilibrium where information is fully transmitted, which in our model is unique. Our focus on this type of equilibrium is important for two reasons. First, we are interested in understanding the negative consequences of countries' private information. A fully revealing equilibrium corresponds to the extreme/worst case in terms of the amount of burned money burned to transmit information. In the other extreme, countries do not incur any costs of money burning, and information can only be transmitted via cheap talk, which may involve interval equilibria as in Crawford and Sobel (1982). Any other semi-separating equilibrium corresponds to an intermediate case between these two extremes. Second, equilibrium uniqueness allows a fair comparison between the two models.

We may also think of a cap on how much money a country can burn (Kartik, 2009). In this case, countries have fewer actions to separate themselves, resulting in a pooling interval for extreme types in equilibrium. From the identical discussion to the previous paragraph, we do not consider this type of equilibrium and instead interpret our equilibrium as one where the cap is sufficiently high.

Formally, fix an equilibrium $(b^{\mathcal{I}}, m^{\mathcal{I}}, d^{\mathcal{I}}, \mu^{\mathcal{I}})$. We illustrate the refinement in the no-delegation game. The refinement for the delegation game follows analogously. Define the following:

$$\underline{\nu}_1(\tilde{b}_1) := \max\left\{\frac{\underline{\theta}_2}{1+\beta} + \frac{\underline{\theta}_1}{1+\beta}\beta, \sup_{\theta_1 : b_1^{ND}(\theta_1) \leq \tilde{b}_1} d_2^{ND}\left(\underline{\theta}_2, b_1^{ND}(\theta_1), m_1^{ND}(\theta_1)\right)\right\},$$

$$\overline{\nu}_1(\tilde{b}_1) := \min\left\{\frac{\overline{\theta}_2}{1+\beta} + \frac{\overline{\theta}_1}{1+\beta}\beta, \inf_{\theta_1 : b_1^{ND}(\theta_1) \geq \tilde{b}_1} d_2^{ND}\left(\overline{\theta}_2, b_1^{ND}(\theta_1), m_1^{ND}(\theta_1)\right)\right\},$$

$$\underline{\nu}_2(\tilde{b}_2) := \max\left\{\frac{\underline{\theta}_1}{1+\beta} + \frac{\underline{\theta}_2}{1+\beta}\beta, \sup_{\theta_2 : b_2^{ND}(\theta_2) \leq \tilde{b}_2} d_1^{ND}\left(\underline{\theta}_1, b_2^{ND}(\theta_2), m_2^{ND}(\theta_2)\right)\right\},$$

$$\overline{\nu}_2(\tilde{b}_2) := \min\left\{\frac{\overline{\theta}_1}{1+\beta} + \frac{\overline{\theta}_2}{1+\beta}\beta, \inf_{\theta_2 : b_2^{ND}(\theta_2) \geq \tilde{b}_2} d_1^{ND}\left(\overline{\theta}_1, b_2^{ND}(\theta_2), m_2^{ND}(\theta_2)\right)\right\}.$$

The function $\underline{\nu}_i(\tilde{b}_i)$ is the lowest equilibrium policy action taken in response to $\tilde{b}_i$. If no type chooses $\tilde{b}_i$, then it is the highest equilibrium policy action taken in response to $b_i \leq \tilde{b}_i$. If no type chooses $b_i \leq \tilde{b}_i$, then it is the highest rationalizable action. The function $\overline{\nu}_i(\tilde{b}_i)$ is the analogue highest equilibrium policy action in response to $\tilde{b}_i$.

Denote

$$\hat{u}_i(d_j, \theta_i, b_i) := \mathbb{E}_i^0\Big[u_i\Big(d_i^{ND}(\theta_i, b_j^{ND}(\theta_j), m_j^{ND}(\theta_j)), d_j, \theta_i, b_i\Big)\Big].$$

Now, define

$$A_i(\tilde{b}_i, \theta_i) := \Big[\underline{\nu}_i(\tilde{b}_i), \overline{\nu}_i(\tilde{b}_i)\Big] \cap \Big\{d_j : \hat{u}_i\Big(d_j, \theta_i, \tilde{b}_i\Big) \geq \hat{u}_i\Big(d_j^{ND}(\theta_j, b_i^{ND}(\theta_i), m_i^{ND}(\theta_i)), \theta_i, b_i^{ND}(\theta_i)\Big)\Big\},$$

$$\overline{A}_i(\tilde{b}_i, \theta_i) := \Big[\underline{\nu}_i(\tilde{b}_i), \overline{\nu}_i(\tilde{b}_i)\Big] \cap \Big\{d_j : \hat{u}_i\Big(d_j, \theta_i, \tilde{b}_i\Big) > \hat{u}_i\Big(d_j^{ND}(\theta_j, b_i^{ND}(\theta_i), m_i^{ND}(\theta_i)), \theta_i, b_i^{ND}(\theta_i)\Big)\Big\}.$$

Fix an amount of burned money $\tilde{b}_i$. The first is the set of responses within the set $[\underline{\nu}_i(\tilde{b}_i), \overline{\nu}_i(\tilde{b}_i)]$ that give type θ_i a weak incentive to deviate to $\tilde{b}_i$. The second one is the strict version of the same set. Let $G_i^{ND}(\cdot|b_j, m_j)$ be the cumulative distribution function of $\mu_i^{ND}(b_j, m_j)$.

An equilibrium $(b^{ND}, m^{ND}, d^{ND}, \mu^{ND})$ satisfies the *mD1 criterion* if it satisfies:

i) b_i^{ND} is a monotonic function.

ii) 1. For all m_1, m_1', θ_1, and $b_1 > b_1'$, $G_2^{ND}(\theta_1|b_1, m_1) \geq G_2^{ND}(\theta_1|b_1', m_1')$.

  2. For all m_2, m_2', θ_2, and $b_2 > b_2'$, $G_1^{ND}(\theta_2|b_2, m_2) \leq G_1^{ND}(\theta_2|b_2', m_2')$.

iii) $Support[\mu_i^{ND}(\tilde{b}_j, \tilde{m}_j)] = \theta_j'$ for any θ_j' and any out-of-equilibrium $\tilde{b}_j$ such that $A_i(\tilde{b}_j, \theta_j) \subseteq \overline{A}_i(\tilde{b}_j, \theta_j')$ for all $\theta_j \neq \theta_j'$ and $A_i(\tilde{b}_j, \theta_j') \neq \emptyset$.

B Proofs for Baseline Model

B.1 Proof of Lemma 0

Following the result from Austen-Smith and Banks (2000), all equilibria have the following structure: there is a partition $(B_0 \equiv \underline{\theta}_i, A_1, B_1, \ldots, A_N, B_N, A_{N+1} \equiv \overline{\theta}_i)$ with $B_{j-1} \leq A_j < B_j \leq A_{j+1}$ for all $j \in I = \{1, \ldots, N\}$, such that a country pools all states $\theta_i \in (A_j, B_j)$ by burning an amount and sending a message $(b_i^{\mathcal{I}}(\theta_i), m_i^{\mathcal{I}}(\theta_i)) = (b_i^{\mathcal{I}}(j), m_i^{\mathcal{I}}(j))$ that is constant over that set. Moreover, if $\theta_i \in (B_j, A_{j+1})$, a country separates by burning different amounts $b_i^{\mathcal{I}}(\theta_i)$. Additionally, for any equilibrium where $\theta_i, \theta_i' \in (B_j, A_{j+1})$ and $m_i^{\mathcal{I}}(\theta_i) \neq m_i^{\mathcal{I}}(\theta_i)$, there is another output equivalent equilibrium where the only difference is that $m_i^{\mathcal{I}}(\theta_i) = m_i^{\mathcal{I}}(\theta_i)$. Thus, it is with out loss of generality we do not need to specify the equilibrium messages for any set of this kind. The partition is uniquely defined by its collection of pooling intervals $P = \{(A_j, B_j)|j \in I\}$. As Austen-Smith and Banks (2000) note, the set of equilibria contains a continuum of semi separating equilibria, with the separating equilibrium $P = \emptyset$ at one end and the pooling equilibrium $P = \{\Theta_i\}$ at the other end.

B.2 Proof of Lemma 1

We sketch a proof that works for both models. Consider the no-delegation game and without loss of generality focus on country 1. Suppose two types $\theta_1 < \theta_1'$ burn the same amount but send different messages m_1, m_1' in equilibrium. It is a profitable deviation for type θ_1' to send message m_1 and pretend to be a lower type to induce a lower policy. Thus, any equilibrium where different types burn the same amount must have these types sending the same message. We show in Proposition 1 and Proposition 2 that b_1^{ND} is a non-increasing function. Intuitively, lower types are willing to burn more because the gains to lie and induce a lower policy are higher. Our refinement rules out cases when types pool on the same amount of burned money.

Suppose by contradiction that b_1^{ND} is not a one-to-one function. There must be an interval $[\theta', \theta''] \subseteq \Theta_1$ such that $b_1^{ND}(\theta) = b^* \geq 0$, $m^{ND}(\theta) = m^*$ for every $\theta \in [\theta', \theta'']$ and $b_1^{ND}(\theta) \neq b^*$ for every $\theta \notin [\theta', \theta'']$. Since the prior belief is uniform, $\mathbb{E}_2[\theta_1|b^*, m^*] = \frac{(\theta''-\theta')}{2}$ and for any $b > b^*$, $\mathbb{E}_2[\theta_1|b, m] \leq \theta'$. Thus

$$\mathbb{E}_2[\theta_1|b, m] \leq \theta' < \frac{(\theta'' - \theta')}{2} = \mathbb{E}_2[\theta_1|b^*, m^*].$$

Invoking the fact that country 1's strategy is sequentially rational, the previous inequality implies

$$\lim_{\theta \to \theta'^-} b_1^{ND}(\theta) > b^*.$$

Consider type $(\theta' - \epsilon)$, and a deviation to an off-path b that satisfies $b^* < b < \lim_{\theta \to \theta'^-} b_1^{ND}(\theta)$. For a sufficiently small ϵ, this type has a profitable deviation to choose b. Intuitively, this deviation is profitable because this type burns a strictly smaller amount b and signals that he is type θ' (because the mD1 criteria restricts the posterior belief after b to put probability one to type θ'), which is ϵ close to his type. This produces a contradiction with type $(\theta' - \epsilon)$'s sequential rationality condition. Then b_1^{ND} is a one-to-one mapping. Thus, the money burned is fully informative of a country's type and characterized by an ODE equation that has a unique solution which is a strictly monotone function (see the proof of Propositions 1 and 2 for more details). Thus, the equilibrium is fully informative and unique. Since information is transmitted with burned money, there is no space for cheap talk messages to produce a different outcome (Austen-Smith and Banks, 2000).

B.3 Proof of Lemma 2

We want to show that in the no-delegation game, country i chooses

$$d_i^{ND} = (1-\beta)\theta_i + \beta \left[\frac{1}{1+\beta} \mathbb{E}_i[\theta_j | b_j, m_j] + \frac{\beta}{1+\beta} \mathbb{E}_j[\theta_i | b_i, m_i] \right].$$

In the final stage, country i has observed θ_i and b_j. For simplicity, denote as $\mathbb{E}_i[\cdot] := \mathbb{E}_i[\cdot | b_j, m_j]$ the expected value with respect to θ_j using country i's beliefs induced by (b_j, m_j). Country i solves

$$\max_{d_i} \mathbb{E}_i \left[-(1-\beta)(d_i - \theta_i)^2 - \beta(d_i - d_j(\theta_j))^2 \right] = \max_{d_i} -(1-\beta)(d_i - \theta_i)^2 - \beta \mathbb{E}_i[d_i - d_j(\theta_j)]^2.$$

Calculating the first-order condition we obtain

$$\begin{aligned} 0 &= -2(1-\beta)(d_i - \theta_i) - 2\beta \mathbb{E}_i(d_i - d_j(\theta_j)) \\ &= -(1-\beta)(d_i - \theta_i) - \beta(d_i - \mathbb{E}_i[d_j(\theta_j)]). \end{aligned}$$

Which results in

$$d_i = (1-\beta)\theta_i + \beta \mathbb{E}_i[d_j(\theta_j)].$$

Analogously, we obtain for country j

$$d_j = (1-\beta)\theta_j + \beta \mathbb{E}_j[d_i(\theta_i)].$$

Taking expected values of each expression

$$\mathbb{E}_i[d_j(\theta_j)] = \mathbb{E}_i[(1-\beta)\theta_j + \beta \mathbb{E}_j[d_i(\theta_i)]] = (1-\beta)\mathbb{E}_i[\theta_j] + \beta \mathbb{E}_j[d_i(\theta_i)],$$

$$\mathbb{E}_j[d_i(\theta_i)] = \mathbb{E}_j[(1-\beta)\theta_i + \beta \mathbb{E}_i[d_j(\theta_j)]] = (1-\beta)\mathbb{E}_j[\theta_i] + \beta \mathbb{E}_i[d_j(\theta_j)].$$

Solving the previous system of equations

$$\mathbb{E}_j[d_i(\theta_i)] = \frac{1}{1+\beta}\mathbb{E}_j[\theta_i] + \frac{\beta}{1+\beta}E_i[\theta_j],$$

$$\mathbb{E}_i[d_j(\theta_j)] = \frac{1}{1+\beta}\mathbb{E}_i[\theta_j] + \frac{\beta}{1+\beta}\mathbb{E}_j[\theta_i].$$

Replacing these expected values, we obtain

$$d_i = (1-\beta)\theta_i + \beta\left[\frac{1}{1+\beta}\mathbb{E}_i[\theta_j] + \frac{\beta}{1+\beta}\mathbb{E}_j[\theta_i]\right],$$

$$d_j = (1-\beta)\theta_j + \beta\left[\frac{1}{1+\beta}\mathbb{E}_j[\theta_i] + \frac{\beta}{1+\beta}\mathbb{E}_i[\theta_j]\right].$$

B.4 Proof of Lemma 3

Let $\mathbb{E}_{IO}[\cdot] := \mathbb{E}_{IO}[\cdot|b_i, m_i, b_j, m_j]$ be the expected value with respect to θ_i and θ_j given the IO's beliefs, which are induced by (b_i, m_i) and (b_j, m_j) respectively. We want to show that in the delegation game the IO chooses

$$d_i^D = \frac{1+\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_i] + \frac{2\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_j].$$

The IO solves

$$\max_{d_i,d_j} \frac{1}{2}\mathbb{E}_{IO}\left[-(1-\beta)((d_i-\theta_i)^2 + (d_j-\theta_j)^2) - \beta((d_i-d_j)^2 + (d_j-d_i)^2)\right].$$

Calculating the first-order condition for d_i we obtain

$$0 = -(1-\beta)(d_i - \mathbb{E}_{IO}[\theta_i]) - 2\beta(d_i - d_j).$$

Analogously, we obtain for d_j

$$0 = -(1-\beta)(d_j - \mathbb{E}_{IO}[\theta_j]) - 2\beta(d_j - d_i).$$

Solving the previous system of equations yields the solutions

$$d_i = \frac{1+\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_i] + \frac{2\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_j],$$
$$d_j = \frac{1+\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_j] + \frac{2\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_i].$$

B.5 Proof of Propositions 1 and 2

Denote $d_i(\theta_i, \theta_i', \theta_j)$ as country i's decision when (i) his type is θ_i, (ii) he signals that his type is θ_i' and (iii) he believes the other country is type θ_j with probability one. Denote analogously country i's political payoff, assuming that country j signals his type truthfully

$$U_i(\theta_i, \theta_i', \theta_j) := \pi_i(d_i(\theta_i, \theta_i', \theta_j), d_j(\theta_j, \theta_j, \theta_i'), \theta_i).$$

Suppose that country i burns $b_i(\theta_i')$ in order to signal that his type is θ_i'. Then, he obtains

$$U_i(\theta_i, \theta_i', \theta_j) - b_i(\theta_i').$$

A function $b_i(\theta_i)$ is incentive-compatible and fully reveals country i's type if

$$\theta_i \in \arg\max_{\theta_i'} \mathbb{E}_i^0 \left[U_i(\theta_i, \theta_i', \theta_j)\right] - b_i(\theta_i'), \text{ and}$$
$$b_i(\theta_i) \text{ is a strictly monotone function.}$$

The first requirement implies that $\theta_i' = \theta_i$ satisfies the following first-order condition

$$\frac{\partial b_i'(\theta_i')}{\partial \theta_i'} = \mathbb{E}_i^0 \left[\frac{\partial U_i(\theta_i, \theta_i', \theta_j)}{\partial \theta_i'}\right].$$

Depending on which institution and country we consider, we have different expressions in the right hand side. Integrating those expressions with respect to θ_i yields the following

- In the case of no delegation:
  i) Country 1 burns $b_1^{ND}(\theta_1) = \frac{2(1-\beta)\beta^2}{(1+\beta)^2}\left(\frac{\theta_1^2}{2} - \theta_1\right) + C$,

  ii) Country 2 burns $b_2^{ND}(\theta_2) = \frac{2(1-\beta)\beta^2}{(1+\beta)^2}\left(\frac{\theta_2^2}{2} + \theta_2\right) + C$.
- In the case of delegation:
  i) Country 1 burns $b_1^{D}(\theta_1) = \frac{2(1-\beta)\beta}{(1+3\beta)}\left(\frac{\theta_1^2}{2} - \theta_1\right) + C$,

  ii) Country 2 burns $b_2^{D}(\theta_2) = \frac{2(1-\beta)\beta}{(1+3\beta)}\left(\frac{\theta_2^2}{2} + \theta_2\right) + C$.

Where C is an integrating constant. These functions are strictly convex centered in 1 for country 1 and in -1 for country 2. For these functions to be equilibrium strategies, the lowest type for each country and in any institution must burn zero. Thus $b_1^{\mathcal{I}}\left(\min\{\overline{\theta}_1, 1\}\right) = b_2^{\mathcal{I}}\left(\max\{\underline{\theta}_2, -1\}\right) = 0$ for any institution $\mathcal{I} \in \{D, ND\}$. Incorporating these restrictions we obtain

- In case of no delegation:
  i) Country 1 burns $b_1^{ND}(\theta_1) = \frac{2(1-\beta)\beta^2}{(1+\beta)^2}\left(\theta_1 - \min\{\overline{\theta}_1, 1\}\right)\left(\frac{\theta_1 + \min\{\overline{\theta}_1, 1\}}{2} - 1\right)$,

  ii) Country 2 burns $b_2^{ND}(\theta_2) = \frac{2(1-\beta)\beta^2}{(1+\beta)^2}\left(\theta_2 - \max\{\underline{\theta}_2, -1\}\right)\left(\frac{\theta_2 + \max\{\underline{\theta}_2, -1\}}{2} + 1\right)$.
- In case of delegation:
  i) Country 1 burns $b_1^{D}(\theta_1) = \frac{2(1-\beta)\beta}{(1+3\beta)}\left(\theta_1 - \min\{\overline{\theta}_1, 1\}\right)\left(\frac{\theta_1 + \min\{\overline{\theta}_1, 1\}}{2} - 1\right)$,

  ii) Country 2 burns $b_2^{D}(\theta_2) = \frac{2(1-\beta)\beta}{(1+3\beta)}\left(\theta_2 - \max\{\underline{\theta}_2, -1\}\right)\left(\frac{\theta_2 + \max\{\underline{\theta}_2, -1\}}{2} + 1\right)$.

If $s \leq 2$, these functions are strictly monotone and a fully revealing equilibrium exists. If $s > 2$, these functions are strictly convex, so for each country, there may be a pair of types that burn the same amount. We can assume that each of these types send a different message to fully reveal their types. For example, if we consider country 1, we can assume types $\theta_1 < 1$ send signal s_l and types $\theta_1 > 1$ send signal s_r, with $s_l \neq s_r$. Thus a fully revealing equilibrium exists too.

Note that $b_i^D(\theta_i) > b_i^{ND}(\theta_i)$, $i \in \{1, 2\}$. Thus for a fixed country and type, the money burned under delegation is higher than the money burned under no delegation.

B.6 Proof of Proposition 3

For any institution $\mathcal{I} \in \{D, ND\}$, denote $\Pi_i^{\mathcal{I}} := \mathbb{E}^0[\pi_i^{\mathcal{I}}(\theta)]$ and $B_i^{\mathcal{I}} := \mathbb{E}^0[b_i^{\mathcal{I}}(\theta_i)]$. In this case countries are symmetric so

$$\Pi^{\mathcal{I}} := \Pi_1^{\mathcal{I}} = \Pi_2^{\mathcal{I}},$$

$$B^{\mathcal{I}} := B_1^{\mathcal{I}} = B_2^{\mathcal{I}}.$$

Using the previous results, after some algebra we obtain

- In case of no delegation:

$$\Pi^{ND} = -\frac{2}{3}\frac{(1-\beta)\beta(6+s^2)}{(1+\beta)^2},$$

$$B^{ND} = \frac{1}{3}\frac{(1-\beta)\beta^2(9+s^2-3\max\{-1,1-s\}^2-6\max\{-1,1-s\})}{(1+\beta)^2}.$$

- In case of delegation:

$$\Pi^{D} = -\frac{2}{3}\frac{(1-\beta)\beta(6+s^2)}{(1+3\beta)},$$

$$B^{D} = \frac{1}{3}\frac{(1-\beta)\beta(9+s^2-3\max\{-1,1-s\}^2-6\max\{-1,1-s\})}{(1+3\beta)}.$$

Note that whenever $\beta \in (0, 1)$, $\Pi^D > \Pi^{ND}$ and $B^D > B^{ND}$.

We need to compare $(\Pi^{ND} - B^{ND})$ and $(\Pi^D - B^D)$. After some algebra we obtain:

$$(\Pi^{ND} - B^{ND}) - (\Pi^D - B^D) = \begin{cases} \frac{2}{3}\frac{(1-\beta)^2\beta\left(s^2+12\beta+12\right)}{(\beta+1)^2(1+3\beta)} & \text{if } s \geq 2 \\ \frac{4}{3}\frac{(1-\beta)^2\beta\left[(6-s)s-3\beta\left(s^2-4s+2\right)\right]}{(1+\beta)^2(1+3\beta)} & \text{if } s < 2. \end{cases}$$

The term $[(6-s)s - 3\beta\,(s^2 - 4s + 2)]$ is increasing in s and has one zero whenever $s < 2$. Let $\hat{s}$ be the value in s that makes zero the last term. Then we have the following:

If $s \leq \hat{s}$, then $(\Pi^{ND} - B^{ND}) \leq (\Pi^D - B^D)$,

If $s > \hat{s}$, then $(\Pi^{ND} - B^{ND}) > (\Pi^D - B^D)$.

Thus, delegation is beneficial only if the level of uncertainty is sufficiently low

$$s \leq \hat{s} := \frac{3 + 6\beta - (9 + 30\beta + 18\beta^2)^{\frac{1}{2}}}{1 + 3\beta}.$$

C Proofs of Extensions

C.1 Proof of Proposition 4

Denote as $U_i^{ND} =: -(1-\beta)(d_i^{ND} - \theta_i)^2 - \beta(d_i^{ND} - d_j^{ND})^2$ the outside option of country i. We first prove the following Lemma.

Lemma 5. *In the equilibrium of the international bargaining game, country $i = 1, 2$ proposes*

$$\begin{aligned}
d_i^{IB} &= \frac{1+\beta}{1+3\beta}\theta_i + \frac{2\beta}{1+3\beta}\mathbb{E}_i[\theta_j] \\
d_j^{IB} &= \frac{1+\beta}{1+3\beta}\mathbb{E}_i[\theta_j] + \frac{2\beta}{1+3\beta}\theta_i \\
T^{IB} &= \mathbb{E}_i\mathbb{E}_j\left[-(1-\beta)(d_j^{IB} - \theta_j)^2 - \beta(d_i^{IB} - d_j^{IB})^2\right] - \mathbb{E}_i\mathbb{E}_j\left[U_j^{ND}\right].
\end{aligned}$$

Country i accepts a proposal (d_i, d_j, T) if and only if

$$\mathbb{E}_i\left[-(1-\beta)(d_i - \theta_i)^2 - \beta(d_i - d_j)^2\right] - T \geq \mathbb{E}_i\left[U_i^{ND}\right]$$

We restrict our analysis to strategies where countries accept an offer when indifferent with probability one. If this is not the case, the maximization problem may not a have solution.

If country i is the proposer, it solves the following problem:

$$\begin{aligned}
\max_{d_i, d_j, T} \quad & \mathbb{E}_i\left[-(1-\beta)(d_i - \theta_i)^2 - \beta(d_i - d_j)^2 + T\right] \\
\text{s.t.} \quad & \mathbb{E}_i\mathbb{E}_j\left[-(1-\beta)(d_j - \theta_j)^2 - \beta(d_i - d_j)^2 - T\right] \geq \mathbb{E}_i\mathbb{E}_j\left[U_j^{ND}\right]
\end{aligned}$$

In the optimum the restriction is binding. The problem becomes:

$$\begin{aligned}
\max_{d_i, d_j} \quad & \mathbb{E}_i\left[-(1-\beta)(d_i - \theta_i)^2 - \beta(d_i - d_j)^2 + \mathbb{E}_j\left[-(1-\beta)(d_j - \theta_j)^2 - \beta(d_i - d_j)^2\right]\right] \\
& - \mathbb{E}_i\mathbb{E}_j\left[U_j^{ND}\right]
\end{aligned}$$

with

$$T = \mathbb{E}_i\mathbb{E}_j\left[-(1-\beta)(d_j - \theta_j)^2 - \beta(d_i - d_j)^2\right] - \mathbb{E}_i\mathbb{E}_j\left[U_j^{ND}\right]$$

The optimum is the following:

$$
\begin{aligned}
d_i^{IB} &= \frac{1+\beta}{1+3\beta}\theta_i + \frac{2\beta}{1+3\beta}\mathbb{E}_i\left[\theta_j\right], \\
d_j^{IB} &= \frac{1+\beta}{1+3\beta}\mathbb{E}_i\left[\theta_j\right] + \frac{2\beta}{1+3\beta}\theta_i, \\
T^{IB} &= \mathbb{E}_i\mathbb{E}_j\left[-(1-\beta)(d_j^{IB}-\theta_j)^2 - \beta(d_i^{IB}-d_j^{IB})^2\right] - \mathbb{E}_i\mathbb{E}_j\left[U_j^{ND}\right].
\end{aligned}
$$

The proposal is accepted because country j is indifferent between the offer and his outside option.

We now restate Proposition 4 formally, with explicit money burning functions:

Proposition 4. *In the equilibrium of the international bargaining game, countries* 1 *and* 2 *burn*

$$
\begin{aligned}
b_1^{IB}(\theta_1) &= \frac{2(1-\beta)\beta\left[(1+\beta)+2((1-p)\beta^2+p\beta)\right]}{(1+\beta)^2(1+3\beta)} f_1(\theta_1), \\
b_2^{IB}(\theta_2) &= \frac{2(1-\beta)\beta\left[(1+\beta)+2(p\beta^2+(1-p)\beta)\right]}{(1+\beta)^2(1+3\beta)} f_2(\theta_2).
\end{aligned}
$$

Moreover, if $p = 1/2$, $b_i^{IB}(\theta_i) = b_i^D(\theta_i)$. *For any value of* p, $b_i^{IB}(\theta_i) > b_i^{ND}(\theta_i)$ *if* $\beta \in (0,1)$. *Finally, since* $\beta > \beta^2$, $b_i^{IB}(\theta_i)$ *is increasing in country* i*'s proposing probability.*

Denote the expected value that country i receives when it proposes as follows:

$$
\begin{aligned}
U_i^i :=& \mathbb{E}_i\left[-(1-\beta)(d_i^{IB}-\theta_i)^2 - \beta(d_i^{IB}-d_j^{IB})^2 + \mathbb{E}_j\left[-(1-\beta)(d_j^{IB}-\theta_j)^2 - \beta(d_i^{IB}-d_j^{IB})^2\right]\right] \\
&- \mathbb{E}_i\mathbb{E}_j\left[U_j^{ND}\right]
\end{aligned}
$$

Denote the expected value that country i receives when country j proposes as follows:

$$
\begin{aligned}
U_i^j :=& \mathbb{E}_i\left[-(1-\beta)(d_i^{IB}-\theta_i)^2 - \beta(d_i^{IB}-d_j^{IB})^2\right] - \mathbb{E}_j\mathbb{E}_i\left[-(1-\beta)(d_i^{IB}-\theta_i)^2 - \beta(d_i^{IB}-d_j^{IB})^2\right] \\
&+ \mathbb{E}_j\mathbb{E}_i\left[U_i^{ND}\right]
\end{aligned}
$$

From an ex-ante perspective, before knowing who is going to be the proposer, country i's payoff is denoted as follows

$$
U^i := pU_i^i + (1-p)U_i^j
$$

Suppose country i is type θ_i, signals his type is θ_i' and he believes the other country is type θ_j with probability one. Denote as $U^i(\theta_i, \theta_i', \theta_j)$ as the ex-ante payoff U^i when the previous is true. Suppose that country i burns $b_i(\theta_i')$ in order to signal his type is θ_i'. Then, he obtains

$$
U^i\left(\theta_i, \theta_i', \theta_j\right) - b_i(\theta_i').
$$

A function $b_i(\theta_i)$ is incentive-compatible and fully reveals country i's type if

$$
\theta_i \in \arg\max_{\theta_i'} \mathbb{E}_i^0\left[U^i(\theta_i, \theta_i', \theta_j)\right] - b_i(\theta_i').
$$

$b_i(\theta_i)$ is a strictly monotone function.

The first requirement implies that $\theta'_i = \theta_i$ satisfies the following first-order condition

$$\frac{\partial b'_i(\theta'_i)}{\partial \theta'_i} = \mathbb{E}^0_i \left[\frac{\partial U^i(\theta_i, \theta'_i, \theta_j)}{\partial \theta'_i} \right].$$

Integrating with respect to θ_i and considering the initial condition gives the following expression for each country

$$b_1^{IB}(\theta_1) = \frac{2(1-\beta)\beta[(1+\beta) + 2((1-p)\beta^2 + p\beta)]}{(1+\beta)^2(1+3\beta)} f_1(\theta_1),$$
$$b_2^{IB}(\theta_2) = \frac{2(1-\beta)\beta[(1+\beta) + 2(p\beta^2 + (1-p)\beta)]}{(1+\beta)^2(1+3\beta)} f_2(\theta_2).$$

C.2 Proof of Proposition 5

We restate Proposition 5 formally:

Proposition 5. *In every equilibrium, the IO proposes*

$$d = \begin{cases} (d_1^{ND} + \frac{\sqrt{c}}{1+\beta}, d_2^{ND} - \frac{\sqrt{c}}{1+\beta}) & \text{if } c \leq \frac{(1-\beta)^2\beta^2\mathbb{E}_{IO}[\theta_1-\theta_2]^2}{(1+3\beta)^2} \\ (d_1^D, d_2^D) & \text{if } c > \frac{(1-\beta)^2\beta^2\mathbb{E}_{IO}[\theta_1-\theta_2]^2}{(1+3\beta)^2} \end{cases}$$

Further, money burning functions b_i^c satisfy the following properties for a fixed θ_i:

i) $b_i^{ND}(\theta_i) \leq b_i^c(\theta_i) \leq b_i^D(\theta_i)$

ii) For $c \leq \underline{c} =: \frac{(1-\beta)^2\beta^2}{(1+3\beta)^2}(2-2s)^2$, $b_i^c(\theta_i) = b_i^{ND}(\theta_i)$

iii) For $c \geq \overline{c} =: \frac{(1-\beta)^2\beta^2}{(1+3\beta)^2}(2+2s)^2$, $b_i^c(\theta_i) = b_i^D(\theta_i)$

iv) $b_i^c(\theta_i)$ *is weakly increasing in* c.

Given an IO's proposal (d_1, d_2), denote as $d_i^{br}(d_j)$ country i's best response policy given d_j, defined as follows:

$$d_i^{br}(d_j) \in \text{argmax} - (1-\beta)(d_i - \theta_i)^2 - \beta(d_i - d_j)^2$$

Note that by construction, $d_i^{br}(d_j^{ND}) = d_i^{ND}$. Moreover, $d_i^{br}(d_j) = (1-\beta)\theta_i + \beta d_j$. Country i will deviate to $d_i^{br}(d_j)$ if:

$$-(1-\beta)(d_i^{br}(d_j) - \theta_i)^2 - \beta(d_i^{br}(d_j) - d_j)^2 - c > -(1-\beta)(d_i - \theta_i)^2 - \beta(d_i - d_j)^2$$

This defines policies for each country for a given IO's recommendation (d_1, d_2).

$$d'_i(d_1, d_2) = \begin{cases} d_i^{br}(d_j) & \text{if country } i \text{ deviates} \\ d_i & \text{if country } i \text{ does not deviate.} \end{cases}$$

In the analysis on which policies the IO will propose, we first look two extreme cases. We then conclude for the intermediate case. If we consider IO's most preferred policies without potential deviations, which corresponds to the solution of the delegation model, we obtains that country i will not deviate if and only if:

$$c \geq \frac{(1-\beta)^2\beta^2}{(1+3\beta)^2}(\theta_i - \theta_j)^2$$

The maximum value that the RHS can take is

$$\bar{c} =: \frac{(1-\beta)^2\beta^2}{(1+3\beta)^2}(\overline{\theta}_2 - \underline{\theta}_1)^2 = \frac{(1-\beta)^2\beta^2}{(1+3\beta)^2}(2+2s)^2 > 0$$

The minimum value that the RHS can take is

$$\underline{c} =: \frac{(1-\beta)^2\beta^2}{(1+3\beta)^2}(\underline{\theta}_2 - \overline{\theta}_1)^2 = \frac{(1-\beta)^2\beta^2}{(1+3\beta)^2}(2-2s)^2 > 0$$

If $c \geq \bar{c}$, both countries will not deviate for any possible types. The IO will propose its most preferred policies as derived in the delegation game, which are going to be accepted

$$d_1 = \frac{1+\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_1] + \frac{2\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_2]$$
$$d_2 = \frac{1+\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_2] + \frac{2\beta}{1+3\beta}\mathbb{E}_{IO}[\theta_1].$$

Suppose $c \leq \underline{c}$. If the IO were to propose its most preferred policies, then both countries would have a profitable deviation. In this case, the IO will optimally propose policies so to make both countires indifferent between the proposed policies and the countries' most profitable deviations.

$$\mathbb{E}_{IO}\left[-(1-\beta)(d_1-\theta_1)^2 - \beta(d_1-d_2)^2\right] = \mathbb{E}_{IO}\left[-(1-\beta)(d_1'(d_2)-\theta_1)^2 - \beta(d_1'(d_2)-d_2)^2\right] - c$$
$$\mathbb{E}_{IO}\left[-(1-\beta)(d_2-\theta_2)^2 - \beta(d_1-d_2)^2\right] = \mathbb{E}_{IO}\left[-(1-\beta)(d_2'(d_1)-\theta_2)^2 - \beta(d_2'(d_1)-d_2)^2\right] - c$$

These give the following policies:

$$d_1 = \frac{1}{1+\beta}\mathbb{E}_{IO}[\theta_1] + \frac{\beta}{1+\beta}\mathbb{E}_{IO}[\theta_2] + \frac{\sqrt{c}}{1+\beta},$$
$$d_2 = \frac{1}{1+\beta}\mathbb{E}_{IO}[\theta_2] + \frac{\beta}{1+\beta}\mathbb{E}_{IO}[\theta_1] - \frac{\sqrt{c}}{1+\beta}$$

Note that $d_1 = d_1^{ND} + \frac{\sqrt{c}}{1+\beta}$ and $d_2 = d_2^{ND} - \frac{\sqrt{c}}{1+\beta}$.

Consider now the incentives to burn money. If $c \geq \bar{c}$, country 1 always anticipates that the IO will propose its most preferred policies. Thus the money burning functions are the same as in the delegation game.

$$b_1(\theta_1) = \frac{2(1-\beta)\beta}{1+3\beta}f_1(\theta_1), \qquad b_2(\theta_2) = \frac{2(1-\beta)\beta}{1+3\beta}f_2(\theta_2).$$

If $c \leq \underline{c}$, country 1 anticipates that the IO will always make both countries indifferent between the proposal and the optimal deviations. Thus the money burning functions are

$$b_1(\theta_1) = \frac{2(1-\beta)\beta^2}{(1+\beta)^2} f_1(\theta_1) - \frac{2(1-\beta)(1+2\beta)}{(1+\beta)^2}\theta_1\sqrt{c} = b_1^{ND}(\theta_1) - \frac{2(1-\beta)(1+2\beta)}{(1+\beta)^2}\theta_1\sqrt{c}$$
$$b_2(\theta_2) = \frac{2(1-\beta)\beta^2}{(1+\beta)^2} f_2(\theta_2) + \frac{2(1-\beta)(1+2\beta)}{(1+\beta)^2}\theta_2\sqrt{c} = b_2^{ND}(\theta_2) + \frac{2(1-\beta)(1+2\beta)}{(1+\beta)^2}\theta_2\sqrt{c}.$$

Note that for a fixed θ_i, b_i is increasing in c.

In general, country 1 would anticipate that the IO proposes its most preferred policies if and only if

$$c \geq \frac{(1-\beta)^2\beta^2}{(1+3\beta)^2}(\theta_2 - \theta_1')^2$$

which is equivalent to

$$\theta_2 < \hat{\theta}_2(\theta_1') := \theta_1' + \frac{(1+3\beta)}{(1-\beta)\beta}\sqrt{c}$$

Denote as $U^i(\theta_i, \theta_i', \theta_j)$ as the ex-ante payoff U^i when the IO proposes its most preferred policies and $U_c^i(\theta_i, \theta_i', \theta_j)$ as the ex-ante payoff U^i when the IO proposes its restricted policies. Define

$$P\left(\theta_j < \hat{\theta}_j(\theta_i')\right) := \frac{\min\{\max\{1+s-\hat{\theta}_j(\theta_i'), 0\}, 1\}}{2s}$$

Thus

$$\begin{aligned}
\mathbb{E}_i^0\left[\frac{\partial U^i(\theta_i,\theta_i',\theta_j)}{\partial \theta_i'}\right] &= \mathbb{E}_i^0\left[\frac{\partial U^i(\theta_i,\theta_i',\theta_j)}{\partial \theta_i'}\middle|\theta_j < \hat{\theta}_j(\theta_i')\right] P\left(\theta_j < \hat{\theta}_j(\theta_i')\right) \\
&+ \mathbb{E}_i^0\left[\frac{\partial U_c^i(\theta_i,\theta_i',\theta_j)}{\partial \theta_i'}\middle|\theta_j > \hat{\theta}_j(\theta_i')\right]\left(1 - P\left(\theta_j < \hat{\theta}_j(\theta_i')\right)\right) \\
&= \int_{1-s}^{\hat{\theta}_j(\theta_i')} \frac{\partial U^i(\theta_i,\theta_i',\theta_j)}{\partial \theta_i'}\frac{1}{2s}d\theta_j \\
&+ \int_{\hat{\theta}_j(\theta_i')}^{1+s} \frac{\partial U_c^i(\theta_i,\theta_i',\theta_j)}{\partial \theta_i'}\frac{1}{2s}d\theta_j.
\end{aligned}$$

Also,

$$\frac{\partial}{\partial c}\mathbb{E}_1^0\left[\frac{\partial U^1(\theta_1,\theta_1',\theta_2)}{\partial \theta_1'}\right] = \int_{\hat{\theta}_2(\theta_1')}^{1+s} \frac{\partial}{\partial c}\frac{\partial U_c^1(\theta_1,\theta_1',\theta_2)}{\partial \theta_1'}\frac{1}{2s}d\theta_2 < 0$$

and

$$\frac{\partial}{\partial c}\mathbb{E}_2^0\left[\frac{\partial U^2(\theta_2,\theta_2',\theta_1)}{\partial \theta_2'}\right] = \int_{\hat{\theta}_1(\theta_2')}^{1+s} \frac{\partial}{\partial c}\frac{\partial U_c^2(\theta_2,\theta_2',\theta_1)}{\partial \theta_2'}\frac{1}{2s}d\theta_1 > 0$$

The money burning function satisfies the following expression when $\theta_i' = \theta_i$:

$$\frac{\partial b_i'(\theta_i')}{\partial \theta_i'} = \mathbb{E}_i^0\left[\frac{\partial U^i(\theta_i,\theta_i',\theta_j)}{\partial \theta_i'}\right].$$

Then, the slope of the function $b_i(\theta_i)$ at θ_i is more pronounced as c increases. Since $b_1(\overline{\theta}_i) = 0$ and $b_2(\underline{\theta}_i) = 0$, for a fixed θ_i the value $b_i(\theta_i)$ is increasing in c.

C.3 Proof of Proposition 6

The proof is analogous to the previous results but we now consider the level of disagreement Δ.

- In case of no delegation:

$$\Pi^{ND} = -\frac{2}{3}\frac{(1-\beta)\beta(s^2+6(\Delta/2)^2)}{(1+\beta)^2}.$$

The money burning functions are

$$\begin{aligned} b_1^{ND}(\theta_1) &= \frac{2(1-\beta)\beta^2}{(1+\beta)^2}\left(\theta_1 - \min\{-\Delta/2+s, \Delta/2\}\right)\left(\frac{(\theta_1 + \min\{-\Delta/2+s, \Delta/2\})}{2} - \frac{\Delta}{2}\right), \\ b_2^{ND}(\theta_2) &= \frac{2(1-\beta)\beta^2}{(1+\beta)^2}\left(\theta_2 - \max\{\Delta/2-s, -\Delta/2\}\right)\left(\frac{\theta_2 + \max\{\Delta/2-s, -\Delta/2\}}{2} + \frac{\Delta}{2}\right). \end{aligned}$$

And finally

$$B^{ND} = \frac{1}{3}\frac{(1-\beta)\beta^2(9(\Delta/2)^2 + s^2 - 3\max\{\Delta/2-s, -\Delta/2\}^2 - 6(\Delta/2)\max\{\Delta/2-s, -\Delta/2\})}{(1+\beta)^2}.$$

- In case of delegation:

$$\Pi^{D} = -\frac{2}{3}\frac{(1-\beta)\beta(s^2+6(\Delta/2)^2)}{(1+3\beta)}.$$

The money burning functions are

$$\begin{aligned} b_1^{D}(\theta_1) &= \frac{2(1-\beta)\beta}{(1+3\beta)}\left(\theta_1 - \min\{-\Delta/2+s, \Delta/2\}\right)\left(\frac{(\theta_1 + \min\{-\Delta/2+s, \Delta/2\})}{2} - \frac{\Delta}{2}\right), \\ b_2^{D}(\theta_2) &= \frac{2(1-\beta)\beta}{(1+3\beta)}\left(\theta_2 - \max\{\Delta/2-s, -\Delta/2\}\right)\left(\frac{\theta_2 + \max\{\Delta/2-s, -\Delta/2\}}{2} + \frac{\Delta}{2}\right). \end{aligned}$$

And finally

$$B^{D} = \frac{1}{3}\frac{(1-\beta)\beta(9(\Delta/2)^2 + s^2 - 3\max\{\Delta/2-s, -\Delta/2\}^2 - 6(\Delta/2)\max\{\Delta/2-s, -\Delta/2\})}{(1+3\beta)}.$$

Note that if $s \leq \Delta$ the money burning functions are strictly monotone so a fully separating equilibrium exists. If $s > \Delta$, a fully revealing equilibrium exists with the help of messages in the same way as Proposition 1 and 2.

Comparing $(\Pi^{ND} - B^{ND})$ and $(\Pi^{D} - B^{D})$ we obtain that there is $\hat{s}_\Delta < \Delta$ such that if $s \leq \hat{s}_\Delta$, then $(\Pi^{ND} - B^{ND}) \leq (\Pi^{D} - B^{D})$ and if $s > \hat{s}_\Delta$, then $(\Pi^{ND} - B^{ND}) > (\Pi^{D} - B^{D})$.

The cutoff $\hat{s}_\Delta$ is the following:

$$\hat{s}_\Delta := \frac{\Delta}{2}\frac{3+6\beta-(9+30\beta+18\beta^2)^{\frac{1}{2}}}{1+3\beta} = \frac{\Delta}{2}\hat{s}.$$

It is direct to check that $\frac{\partial \hat{s}_\Delta}{\partial \Delta} = \frac{\hat{s}}{2} > 0$.

C.4 Proof of Proposition 7

Now assume country 1's type is publicly observable. Since country 1 can not influence beliefs through its signals, it does not burn money. Political payoffs are the same as in the previous results.

No-delegation. The money burning functions are the following

$$\begin{aligned} b_1^{ND}(\theta_1) &= 0, \\ b_2^{ND}(\theta_2) &= \frac{2(1-\beta)\beta^2}{(1+\beta)^2}(\theta_2 - \max\{1-s,\theta_1\})\left(\frac{\theta_2+\max\{1-s,\theta_1\}}{2} - \theta_1\right). \end{aligned}$$

And then

$$B_2^{ND} = \begin{cases} \frac{2}{3}\frac{(1-\beta)\beta^2(s^3+6s-2)}{(1+\beta)^2 s} & \text{if } s \geq 1 \\ \frac{2}{3}\frac{(1-\beta)\beta^2(6-s)s}{(1+\beta)^2} & \text{if } s < 1, \end{cases}$$

Delegation. The money burning functions are the following

$$\begin{aligned} b_1^{D}(\theta_1) &= 0, \\ b_2^{D}(\theta_2) &= \frac{2(1-\beta)\beta}{(1+3\beta)}(\theta_2 - \max\{1-s,\theta_1\})\left(\frac{\theta_2+\max\{1-s,\theta_1\}}{2} - \theta_1\right). \end{aligned}$$

And then

$$B_2^{D} = \begin{cases} \frac{2}{3}\frac{(1-\beta)\beta(s^3+6s-2)}{(1+3\beta) s} & \text{if } s \geq 1 \\ \frac{2}{3}\frac{(1-\beta)\beta(6-s)s}{(1+3\beta)} & \text{if } s < 1. \end{cases}$$

We have that $B_2^{ND} < B_2^D$. Comparing terms, we obtain that $\Pi_1^D > \Pi_1^{ND}$, thus country 1 always prefers to delegate. In the other side, after some algebra we obtain

If $s \leq \hat{s}$, then $(\Pi_2^{ND} - B_2^{ND}) \leq (\Pi_2^D - B_2^D)$,

If $s > \hat{s}$, then $(\Pi_2^{ND} - B_2^{ND}) > (\Pi_2^D - B_2^D)$.

The cutoff is the same than proposition 3. In case we consider $(\Pi_1^{ND} + \Pi_2^{ND} - B_2^{ND}) - (\Pi_1^D + \Pi_2^D - B_2^D)$, there is $\check{s}$ with $\check{s} > \hat{s}$ such that:

If $s \leq \check{s}$, then $(\Pi_1^{ND} + \Pi_2^{ND} - B_2^{ND}) \leq (\Pi_1^D + \Pi_2^D - B_2^D)$,

If $s > \check{s}$, then $(\Pi_1^{ND} + \Pi_2^{ND} - B_2^{ND}) > (\Pi_1^D + \Pi_2^D - B_2^D)$.

$$s < \check{s} := \frac{\left(3+6\beta-(9+24\beta-12\beta^2)^{\frac{1}{2}}\right)}{(1+4\beta)}.$$

C.5 Proof of Lemma 4

We calculate equilibrium payoffs as a function of α and then we study the case $\alpha = 1/2$. Now, policy payoffs are as follows:

$$\begin{aligned}
\pi_1(d_1, d_2, \theta_1) &= -(d_1 - \theta_1)^2, \\
\pi_2(d_2, d_1, \theta_2) &= -(1 - \beta_2)(d_2 - \theta_2)^2 - \beta_2(d_2 - d_1)^2.
\end{aligned}$$

In the case of delegation the IO maximizes the following

$$u_{IO}(d_1, d_2, \theta_1, \theta_2) = \alpha \left[-(d_1 - \theta_1)^2\right] + (1 - \alpha) \left[-(1 - \beta_2)(d_2 - \theta_2)^2 - \beta_2(d_2 - d_1)^2\right].$$

We need to obtain ex ante political and informational payoffs for both cases.

- In case of no delegation:

Countries take the following decisions as a function of θ_1 and θ_2

$$\begin{aligned}
d_1^{ND} &= \theta_1, \\
d_2^{ND} &= \beta_2\theta_1 + (1 - \beta_2)\theta_2.
\end{aligned}$$

Thus, ex ante political payoffs are as follows

$$\begin{aligned}
\Pi_1^{ND} &= 0, \\
\Pi_2^{ND} &= -\frac{2}{3}(1 - \beta_2)\beta_2(6 + s^2).
\end{aligned}$$

Finally, countries have no incentives to burn money since country 1 does not benefit from manipulation and country 2 can not influence. Thus

$$B_1^{ND} = B_2^{ND} = 0.$$

- In case of delegation:

The IO chooses the following decisions as a function of θ_1 and θ_2:

$$\begin{aligned}
d_1^D &= \frac{\theta_1\alpha + \theta_2(1 - \alpha)(1 - \beta_2)\beta_2}{\alpha + \beta_2 - \alpha\beta_2 - \beta_2^2 + \alpha\beta_2^2}, \\
d_2^D &= \frac{\theta_1\alpha\beta_2 + \theta_2(1 - \beta_2)(\beta_2 + \alpha(1 - \beta_2))}{\alpha + \beta_2 - \alpha\beta_2 - \beta_2^2 + \alpha\beta_2^2}.
\end{aligned}$$

Ex ante political payoffs are as follows

$$\begin{aligned}
\Pi_1^D &= -\frac{2}{3}\frac{(1 - \alpha)^2(1 - \beta_2)^2\beta_2^2(6 + s^2)}{(\alpha + \beta_2 - \alpha\beta_2 - (1 - \alpha)\beta_2^2)^2}, \\
\Pi_2^D &= -\frac{2}{3}\frac{\alpha^2(1 - \beta_2)\beta_2(6 + s^2)}{(\alpha + \beta_2 - \alpha\beta_2 - (1 - \alpha)\beta_2^2)^2}.
\end{aligned}$$

Money burning functions

$$b_1^D(\theta_1) = \frac{2(1 - \alpha)\alpha(1 - \beta_2)\beta_2}{(\alpha + \beta_2 - \alpha\beta_2 - \beta_2^2 + \alpha\beta_2^2)^2}\left(\theta_1 - \min\{\overline{\theta}_1, 1\}\right)\left(\frac{\theta_1 + \min\{\overline{\theta}_1, 1\}}{2} - 1\right),$$

$$b_2^D(\theta_2) = \frac{2(1-\alpha)\alpha(1-\beta_2)^2\beta_2^2}{(\alpha+\beta_2-\alpha\beta_2-\beta_2^2+\alpha\beta_2^2)^2}\left(\theta_2 - \max\{\underline{\theta}_2, -1\}\right)\left(\frac{\theta_2+\max\{\underline{\theta}_2,-1\}}{2}+1\right).$$

Then, ex ante informational payoff

$$B_1^D = \frac{2}{3}\frac{(1-\alpha)\alpha(1-\beta_2)\beta_2(6-s)s}{(\alpha+\beta_2-\alpha\beta_2-(1-\alpha)\beta_2^2)^2},$$

$$B_2^D = \frac{2}{3}\frac{(1-\alpha)\alpha(1-\beta_2)^2\beta_2^2(6-s)s}{(\alpha+\beta_2-\alpha\beta_2-(1-\alpha)\beta_2^2)^2}.$$

The rest of the proof assume $\alpha = 1/2$. After some algebra, we obtain $(\Pi_1^{ND} - B_1^{ND}) > (\Pi_1^D - B_1^D)$ and $(\Pi_2^{ND} - B_2^{ND}) < (\Pi_2^D - B_2^D)$. Thus country 1 prefers not to delegate while country 2 prefers to delegate. If we consider instead

$$(\Pi_1^{ND} - B_1^{ND} + \Pi_2^{ND} - B_2^{ND}) - (\Pi_1^D - B_1^D + \Pi_2^D - B_2^D),$$

We obtain that there is $\tilde{s}$ such that

If $s \leq \tilde{s}$, then $(\Pi_1^{ND} - B_1^{ND} + \Pi_2^{ND} - B_2^{ND}) \leq (\Pi_1^D - B_1^D + \Pi_2^D - B_2^D)$,

If $s > \tilde{s}$, then $(\Pi_1^{ND} - B_1^{ND} + \Pi_2^{ND} - B_2^{ND}) > (\Pi_1^D - B_1^D + \Pi_2^D - B_2^D)$.

The cutoff $\tilde{s}$ is the following

$$\tilde{s} := \frac{3 - (9 - 6\beta_2 + 12\beta_2^3 - 6\beta_2^4)^{\frac{1}{2}}}{1+\beta_2-\beta_2^2}.$$

C.6 Proof of Proposition 8

We now use the previous results and study the case of general α. Consider $(\Pi_1^D - B_1^D + \Pi_2^D - B_2^D)$ as a function of α. Define

$$\alpha(\beta, s) := \frac{-(1-\beta)\beta(s^2(\beta^2-\beta-3)+s(-6\beta^2+6\beta+6)-12)}{s(s-6)-2\beta^3 s(s-6)+\beta^4 s(s-6)-\beta^2(5s^2-6s+24)+6\beta(s^2-2s+4)},$$

and

$$s^* := \frac{3+3\beta_2-3\beta_2^2-(9+6\beta_2-33\beta_2^2+54\beta_2^3-27\beta_2^4)^{1/2}}{1+3\beta_2-3\beta_2^2}.$$

Denote $\hat{\alpha}$ the maximizer of $(\Pi_1^D - B_1^D + \Pi_2^D - B_2^D)$ restricted to $0 \leq \alpha \leq 1$. A simple first-order condition analysis implies the following

If $0 < s \leq s^*$, then $\hat{\alpha} = \alpha(\beta_2, s)$,

If $s^* < s$, then $\hat{\alpha} = 1$.

It is direct to see that $\frac{\partial \alpha(\beta_2,s)}{\partial s} > 0$ and $\alpha(\beta_2, 0) = \frac{1}{2}$.

C.7 Proof of Proposition 9

- In the case of no delegation, the expressions are the same as in Proposition 4.
- In the case of delegation:

We impose the restriction that $\ell \in [0, 2s]$, because when $\ell > 2s$, policies are the same as with $\ell = 2s$. The reason is that the highest and lowest policy the IO ever takes are

$$\begin{aligned} d^{max} &= \frac{\max\theta_1 + \max\theta_2}{2} = \frac{-\Delta + s + \Delta + s}{2}, \\ d^{min} &= \frac{\min\theta_1 + \min\theta_2}{2} = \frac{-\Delta - s + \Delta - s}{2}. \end{aligned}$$

The difference between the two is the set of policies that the IO will possibly take in equilibrium, which equals $\ell = d^{max} - d^{min} = 2s$. Further, taking expected values we obtain the following ex ante political payoff for both countries:

$$\Pi^D = -\frac{1}{3}(1-\beta)\left(\frac{3}{4}\left(\ell^2 + \Delta^2\right) - \ell s + s^2\right).$$

The money burning functions are the following

$$\begin{aligned} b_1^D(\theta_1) &= \frac{(1-\beta)\ell}{2s}\left[\theta_1^2 - \min\{-\Delta/2 + s, 0\}^2\right], \\ b_2^D(\theta_2) &= \frac{(1-\beta)\ell}{2s}\left[\theta_2^2 - \max\{\Delta/2 - s, 0\}^2\right]. \end{aligned}$$

Taking an expectation leads to the following ex ante informational payoff for each country:

$$B^D = \frac{\ell}{6s}(1-\beta)\left(s^2 + 3(\Delta/2)^2 - 3\min\{-\Delta/2 + s, 0\}^2\right).$$

Consider the function $(\Pi^D - B^D)$. If $\ell \geq 2s$, then $(\Pi^D - B^D) < (\Pi^{ND} - B^{ND})$. If we optimize the expression $(\Pi^D - B^D)$ restricted to $\ell \in [0, 2s]$ we obtain:

$$\ell(s, \Delta) = \begin{cases} 0 & \text{if } 0 \leq s \leq \sqrt{3}(\Delta/2) \\ \frac{s}{3} - \frac{\Delta^2}{4s} & \text{if } \sqrt{3}(\Delta/2) < s \leq \Delta. \end{cases}$$

C.8 Proof of Proposition 10

In this extension, we suppose that the IO proposes (d_1, d_2, T), which has to be accepted from both countries. In case it is rejected, we assume countries play the non-delegation game. We let the IO's payoff be the following:

$$u_{IO}(d_1, d_2, \theta_1, \theta_2) = -\alpha(d_1 - d_2)^2.$$

The parameter $\alpha > 0$ measures the IO's coordination motive. We study how this affect countries' signaling incentives.

Proposition 10. *In equilibrium, $b_i^\alpha(\theta_i)$ is increasing in the IO's coordination motive α.*

Denote as $U_i^{ND} =: -(1-\beta)(d_i^{ND} - \theta_i)^2 - \beta(d_i^{ND} - d_j^{ND})^2$ the outside option of country

i. The IO solves

$$\max_{d_i,d_j,T_i,T_j} \quad \mathbb{E}_{IO}\left[-\alpha(d_i-d_j)^2\right]+T_i+T_j$$

s.t.

$$\mathbb{E}_{IO}\mathbb{E}_i\left[-(1-\beta)(d_i-\theta_i)^2-\beta(d_i-d_j)^2\right]-T_i \geq \mathbb{E}_{IO}\mathbb{E}_i\left[U_i^{ND}\right]$$
$$\mathbb{E}_{IO}\mathbb{E}_j\left[-(1-\beta)(d_j-\theta_j)^2-\beta(d_i-d_j)^2\right]-T_j \geq \mathbb{E}_{IO}\mathbb{E}_j\left[U_j^{ND}\right]$$

In the optimum the restrictions are binding. The problem becomes:

$$\begin{aligned}\max_{d_i,d_j} \quad & \mathbb{E}_{IO}\left[-(1-\beta)((d_i-\theta_i)^2+(d_j-\theta_j)^2)-(\alpha+2\beta)(d_i-d_j)^2\right] \\ & -\mathbb{E}_{IO}\left[\mathbb{E}_i\left[U_i^{ND}\right]+\mathbb{E}_j\left[U_j^{ND}\right]\right]\end{aligned}$$

with

$$\begin{aligned}T_i &= \mathbb{E}_{IO}\mathbb{E}_i\left[-(1-\beta)(d_i-\theta_i)^2-\beta(d_i-d_j)^2\right]-\mathbb{E}_{IO}\mathbb{E}_i\left[U_i^{ND}\right] \\ T_j &= \mathbb{E}_{IO}\mathbb{E}_j\left[-(1-\beta)(d_j-\theta_j)^2-\beta(d_i-d_j)^2\right]-\mathbb{E}_{IO}\mathbb{E}_j\left[U_j^{ND}\right]\end{aligned}$$

The optimum is the following:

$$\begin{aligned}d_i^{IO} &= \frac{1+\alpha+\beta}{1+2\alpha+3\beta}\mathbb{E}_{IO}\left[\theta_i\right]+\frac{\alpha+2\beta}{1+2\alpha+3\beta}\mathbb{E}_{IO}\left[\theta_j\right], \\ d_j^{IO} &= \frac{1+\alpha}{1+2\alpha+3\beta}\mathbb{E}_{IO}\left[\theta_j\right]+\frac{\alpha+2\beta}{1+2\alpha+3\beta}\mathbb{E}_{IO}\left[\theta_i\right] \\ T_i^{IO} &= \mathbb{E}_{IO}\mathbb{E}_i\left[-(1-\beta)(d_i^{IO}-\theta_i)^2-\beta(d_i^{IO}-d_j^{IO})^2\right]-\mathbb{E}_{IO}\mathbb{E}_i\left[U_i^{ND}\right] \\ T_j^{IO} &= \mathbb{E}_{IO}\mathbb{E}_j\left[-(1-\beta)(d_j^{IO}-\theta_j)^2-\beta(d_i^{IO}-d_j^{IO})^2\right]-\mathbb{E}_{IO}\mathbb{E}_j\left[U_j^{ND}\right]\end{aligned}$$

The proposal is accepted by each country since both are indifferent between the offer and outside option. In this case

$$b_i^{\alpha}(\theta_i)=\frac{2(1-\beta)(\alpha+\beta+\beta^2+\alpha\beta^2+2\beta^3)}{(1+\beta)^2(1+2\alpha+3\beta)}f_i(\theta_i)$$

Note that $b_i^{\alpha}(\theta_i)$ is increasing in α. Also $b_i^{\alpha}(\theta_i)>b_i^{ND}(\theta_i)$.