Global Public Goods Provision, Information Dissemination, and Domestic Politics[*]

Justin Melnick[†]

March 11, 2025

Abstract

International institutions increasingly adopt flexible arrangements based on voluntary commitments and transparency mechanisms to enhance global public goods provision. These frameworks, like pledge-and-review mechanisms, aim to foster cooperation by publicizing national commitments rather than enforcing compliance through penalties. Conventional wisdom suggests that flexibility and transparency improve cooperation by lowering institutional constraints and enhancing information dissemination. However, this paper demonstrates that such mechanisms can have unintended consequences by weakening leaders' domestic incentives to contribute. Using a formal model, I show that transparency in international agreements disrupts domestic political dynamics by reducing leaders' incentives to signal competence through public goods investments. By making leaders' commitments publicly available, these agreements undermine accountability relationships, weakening incentives to exert costly effort. Paradoxically, international transparency may reduce cooperation rather than enhance it. This study highlights a tradeoff between international cooperation and domestic political survival, offering new insights into institutional design and the effectiveness of pledge-and-review frameworks.

International institutions have developed increasingly flexible institutional arrangements focused on voluntary national commitments and international transparency mechanisms to facilitate global public goods provision. These institutional features, like pledge-and-review frameworks (Harstad 2023a;b), allow political leaders to determine the extent to which they wish to contribute to global provision efforts; institutions aggregate and disseminate information about each state's willingness to cooperate rather than impose pecuniary penalties for noncompliance (see Harrison and Lagunoff 2017; Slechten 2020; McAllister and Schnakenberg 2022, in the context of international climate agreements). Can these institutions succeed in inspiring ambitious contributions?[1]

This note studies a formal model in which political leaders, who are accountable to domestic audiences, are party to an international organization in which they voluntarily reveal their willingness to invest in global public goods and subsequently exert costly effort into contributions. The organization makes leaders' reports public. Conventional wisdom on the structure of international agreements would contend that these features should increase cooperation due to increased flexibility (Rosendorff and Milner 2001) and transparency (Dai 2002) and that the dissemination of information improves cooperation (Keohane 1984; Dai 2005). Contrarily, I demonstrate that these features can have deleterious effects on leaders' investments into public goods because the agreement's design inspires less ambitious commitments by nullifying *domestic* incentives to contribute.

I highlight a novel theoretical tradeoff between international cooperation and leaders' domestic political survival. While institutions that seek to elicit leaders' private information intend to *screen* leaders based on their willingness to contribute to global public goods, domestic accountability relationships may incentivize leaders to *signal* competence to audiences at home through policy implementation. If institutions publicize reports about leaders' willingness to contribute to public goods, they undermine any signaling incentives that may encourage greater investments.

To study how an international agreement's informational effects alter the prospects for cooperation, I contrast equilibria of two settings. First I study the interaction between a leader and a voter without the agreement. This interaction is a simple selection and control relationship (cf.

---

[1]Examples include the Paris Climate Accords, the OECD's Development Assistance Committee Peer Reviews, the UN Convention Against Corruption's Implementation Review Mechanism, the World Trade Organization's Trade Policy Review Mechanism, the Universal Periodic Review of the Human Rights Council, and the International Labor Organization's Complaint Procedure (Raiser, Çalı and Flachsland 2022).

Ashworth, Bueno de Mesquita and Friedenberg 2017): leaders vary in their marginal costs of exerting effort into investments (competence), which determines willingness to make contributions. The voter observes an imperfect signal of effort, the implemented policy, informative of the leader's competence. Reelection prospects are conditioned on observable policy "success" of public goods provision; leaders find it politically advantageous to exert effort because it signals competence.

The second model introduces the international organization. Leaders make a report of their type to the organization—representing their intended contribution to the public good—which the organization publicizes. The core insight is that the agreement's transparency mechanisms provide information about leader type to voters, which weakens the accountability relationship by resolving uncertainty about the leader's competence (cf. Prat 2005; Fox and Jordan 2011). Consequently, voter assessments of a leader's willingness to invest in public goods are no longer conditioned on leader effort, hence leaders' electoral incentives to exert effort dissipate. Paradoxically, with an agreement in place, the provision of public goods is expected to be lower than without it.

This paper contributes to literatures on institutional design, particularly those studying the role of information in international cooperation. I identify a caveat in the conventional wisdom that international organizations' revelation of information can improve cooperation and quality of governance (Keohane 1984; Kelley and Simmons 2019). If cooperative policies have domestic ramifications, leaders may be inspired to exert greater effort due to the domestic pressures to deliver rather than from threats of international sanctioning; hence, the design of agreements may have unintended negative consequences for international cooperative goals. Moreover, I contribute to theoretical literature on pledge-and-review frameworks (Harstad 2023a) by studying the domestic incentives within these arrangements and how they interact with institutional constraints.

Model Setup

Consider a strategic interaction between $n > 2$ countries that engage in policymaking on the provision of a global public good. I contrast two cases: in the first, countries act independently and in the second they are engaged with an international organization (IO).

In each country (indexed by i) there is a leader and a representative voter. Leader i has a

private type, $\theta_i \in \{\underline{\theta} \ \overline{\theta}\}$ with $0 < \underline{\theta} < \overline{\theta}$. Let the common prior be $P(\theta_i = \underline{\theta}) = q$. Only leader i knows her type; all voters and all other leaders only know the prior. Types are the ease with which leaders produce public goods, their competence. Think of competence as a marginal cost: leaders of type $\underline{\theta}$ have lower costs of producing public goods and are all else equal more willing to invest.[2]

Given their types, leaders allocate effort toward public good provision, $a_i \in [0, 1]$. Let leader i's utility for public goods be $u(a; \theta_i)$, which is assumed to be twice continuously differentiable in a_i, increasing and strictly concave. Let the cross-partial of effort a_i and type θ_i satisfy $\frac{\partial^2 u(a;\theta_i)}{\partial a_i \partial \theta_i} \leq 0$, formalizing θ_i as leader i's marginal cost of producing the public good. Finally, I assume that $u(a; \theta_i)$ is additively separable between its own effort and the effort exerted by other countries. This eases exposition and allows me to concentrate on the effects of information that persist while holding fixed any collective action concerns. To fix ideas, an example satisfying these criteria is

$$u(a; \theta_i) = \sum_{i=1}^{n} a_i - \frac{\theta_i}{2} a_i^2.$$

Each leader is better off when nations work harder toward providing global public goods, as leader i's utility increases in the efforts of other nations; however, effort is individually costly, captured by the quadratic term.

Let the level of effort that satisfies $\frac{\partial u(a;\theta_i)}{\partial a_i} = 0$ be leader i's "ideal point" level and denote it as $\tilde{a}_i$. This is the level of effort that maximizes $u(a; \theta_i)$ and serves as a convenient benchmark because it represents the level of effort that leaders would exert if the accountability mechanism described below was turned off. With the utility function above, $\tilde{a}_i = \frac{1}{\theta_i}$.

Within the agreement, the IO solicits information about each leader's type θ_i, which is done in the form of a report, and recommends efforts based on the reports. Reports and recommendations are then made public to all other players. To characterize the optimal effort recommendations, I utilize the tools of mechanism design (cf. McAllister and Schnakenberg 2022). Let leader i's report be $\hat{\theta}_i \in \{\underline{\theta}, \overline{\theta}\}$, with the IO's corresponding recommendation being $x(\hat{\theta}_i)$ (which is a recommended

---

[2]Investing in global public goods may have political consequences for leaders because these policies could create domestic winners and losers. Alternatively, leaders vary in their *ex ante* ideological proclivity toward investing in international cooperative projects.

level of a_i). The IO designs $x(\hat{\theta}_i)$ to maximize

$$V = \sum_i u(x(\hat{\theta}); \theta_i).$$

Applying the revelation principle, I examine the class of direct mechanisms that the IO could design such that its effort recommendations would be implementable given the leader's interaction with the voter (more on voters below). As is standard, the IO considers leaders' information and strategic constraints. The information constraint requires that it is incentive compatible for leader i to reveal its true type θ_i, rather than lie and report type θ_i', given that other leaders also submit truthful reports. The strategic constraint requires that it must be weakly profitable to obey the IO's recommended level of effort so that when party to the agreement, leader i would choose $a_i = x(\hat{\theta}_i)$.

Leaders are also accountable to domestic publics when investing in global public goods. The voter sees a noisy signal of leader i's effort, $K_i = a_i + \varepsilon_i$, where $\varepsilon_i \sim N(0,1)$. The signal literally implies that voters have imperfect information about leader effort, but could also represent the policy's "success." Based on the realization of the signal, voter i determines whether to retain leader i or replace her, $\rho_i \in \{0,1\}$ where $\rho_i = 1$ denotes retention and $\rho_i = 0$ denotes replacement.

The voter values competent leaders, getting a payoff normalized to 1 for having a competent leader in office.[3] He also values the incumbent on all other electorally relevant dimensions besides the implementation of public goods, represented by the valence shock $y_i \sim U[-\gamma, \gamma]$.[4][5] The parameter γ governs the salience of public goods provision relative to all other issues, a large γ implies public goods provision is a less salient issue for the voter. The value of this shock is realized right before the voter makes his choice to retain the leader or not. If the voter replaces the incumbent leader, her replacement is of type θ_C such that $P(\theta_C = \underline{\theta}) = q$. The voter's utility function is thus:

$$v(\rho_i; y_i) = \rho_i \left(\mathbb{1}_{\theta_i = \underline{\theta}} + y_i \right) + (1 - \rho_i) \mathbb{1}_{\theta_C = \underline{\theta}}.$$

---

[3]This is a reduced form of a two-period model where voter utility is $v = K_1 + K_2$. There is a selection problem in favor of retaining competent types because they produce public goods more cheaply.

[4]This shock is not formally necessary but smooths out reelection probabilities in the analysis.

[5]I assume the support of the shock γ is sufficiently large, a sufficient condition to ensure that the incompetent leader's utility is concave. See proof of Proposition 1 in the appendix.

Leaders value public goods provision $u(a; \theta_i)$ and an officeholding benefit $\Psi > 0$ if retained:

$$U(a, \rho_i; \theta_i) = u(a; \theta_i) + \rho_i \Psi.$$

The exercise at hand is to determine the IO's optimal recommendations $x(\underline{\theta})$ and $x(\overline{\theta})$ in the shadow of domestic electoral selection and contrast them with leaders' equilibrium effort choices $a^*(\theta_i)$ without the agreement. Timing is as follows:

1. Types θ_i are revealed to leaders. Leaders simultaneously submit reports of type $\hat{\theta}_i$ to the IO.

2. The IO recommends effort $x(\hat{\theta}_i)$ to leaders. Reports and recommendations are made public. Leaders choose effort $a(\theta_i)$.

3. Voters observe the signal K_i and valence shock y_i and retain or replace their leaders.

I look for weak Perfect Bayesian equilibria. Without the agreement, a leader's strategy is an effort level given type. In the agreement, it is a report of type that maps onto a recommended effort level. A strategy for the voter is a retention rule given posterior beliefs about the leader's type. See the appendix for further comments on model assumptions.

Analysis

Optimal Efforts with No Agreement

Without the agreement, the interaction between the leader and the voter is straightforward. Observing policy implementation K_i, the voter forms a posterior belief that the leader is competent $\mu(K_i) = P(\theta_i = \underline{\theta} | K_i)$ and adopts a cutoff rule that satisfies $\mu(K_i) + y_i = q$, retaining the leader if $K_i \geq \hat{K}(y_i)$—when the implementation of public goods is sufficiently successful—and replacing her otherwise. Given this cutoff, the leader knows her expected probability of reelection, and her

effort choice $a^*(\theta_i)$ satisfies the following first-order condition:

$$\underbrace{\frac{\partial u(a;\theta_i)}{\partial a_i}\Big|_{a_i=a^*(\theta_i)}}_{\substack{\text{marginal willingness to} \\ \text{exert effort}}} + \underbrace{\frac{\Psi}{2\gamma}\int_{-1+q}^{q}\phi\Big(a^*(\theta_i)-\hat{K}^*(y)\Big)\,dy}_{\substack{\text{marginal benefit of} \\ \text{office-holding}}} = 0.$$

Leaders balance the marginal costs of exerting effort with their electoral returns and the marginal policy benefits from contribution. With electoral incentives, leaders always behave more ambitiously compared to their ideal point, because doing so signals competence to the voter, $a^*(\theta_i) > \tilde{a}(\theta_i)$. Also, competent leaders invest more than incompetent leaders because they can do so more cheaply and this enhances their electoral prospects. The following proposition highlights relevant features of the equilibrium (with proofs of all formal results in the appendix).

Proposition 1 *In the unique PBE without the international agreement:*

- *if $\Psi = 0$, leaders exert effort at their ideal points, $a^*(\theta_i) = \tilde{a}(\theta_i)$;*

- *effort increases in the value of office-holding, $\frac{\partial}{\partial \Psi}a^*(\theta_i) \geq 0$;*

- *competent leaders exert greater effort than incompetent leaders, $a^*(\underline{\theta}) > a^*(\overline{\theta})$.*

Because leaders are responsive to electoral incentives ($\Psi > 0$), they are motivated to signal competence by exerting greater effort into public goods provision. Their accountability relationship with the voter motivates greater effort than they would ideally prefer, $a^*(\theta_i) > \tilde{a}(\theta_i)$.

Optimal Effort in the Agreement

Now consider the case where leaders report their willingness to contribute to the IO, which gives effort recommendations. I am interested in finding the recommendations that could be made where leaders report their type truthfully (incentive compatibility) and would be followed (obedience).

I proceed first by considering the scope of incentive compatible recommendations the IO could prescribe (see the approach by Harrison and Lagunoff 2017; McAllister and Schnakenberg 2022). Incentive constrains are an important screening mechanism as the IO would like competent leaders

to exert more effort, and they can do so more cheaply. This engenders a temptation for leaders to overstate their contribution costs to receive more lenient recommendations.

An incentive compatible recommendation induces truthful revelation of leader type. It must be optimal for leader of type θ_i to report $\hat{\theta}_i = \theta_i$ rather than lie and report $\hat{\theta}_i = \theta'_i$. This is important for domestic politics, as it means that voters learn the types of their leaders with certainty through the IO's reports. Immediately, this means the voter's posterior belief μ about leader competence no longer depends on the signal K_i, and, subsequently, neither do leaders' reelection prospects. The voter retains the leader if $y_i \geq q - \mu$ which occurs with probability $\frac{\gamma - q + \mu}{2\gamma}$ (where μ is either 0 or 1). Incentive compatibility constraints for each leader type are thus:

$$\underbrace{u(x(\underline{\theta}); \underline{\theta})}_{\substack{\text{policy payoff for} \\ \text{truthfully reporting } \theta_i = \underline{\theta}}} + \underbrace{\frac{1}{2\gamma}\Psi}_{\substack{\text{net electoral benefit for} \\ \text{reporting } \theta_i = \underline{\theta}}} \geq \underbrace{u(x(\overline{\theta}); \underline{\theta})}_{\substack{\text{policy payoff for} \\ \text{reporting } \theta_i = \overline{\theta}}}.$$

$$\underbrace{u(x(\overline{\theta}); \overline{\theta})}_{\substack{\text{policy payoff for} \\ \text{truthfully reporting } \theta_i = \overline{\theta}}} \geq \underbrace{u(x(\underline{\theta}); \overline{\theta})}_{\substack{\text{policy payoff for} \\ \text{reporting } \theta_i = \underline{\theta}}} + \underbrace{\frac{1}{2\gamma}\Psi}_{\substack{\text{net electoral benefit for} \\ \text{reporting } \theta_i = \underline{\theta}}}.$$

In any incentive compatible reporting scheme, incompetent leaders are immediately less likely to survive in office than competent leaders, as $\frac{\gamma - q}{2\gamma} < \frac{\gamma - q + 1}{2\gamma}$. To compensate, the IO must require that competent leaders must always invest more effort than incompetent leaders, $x^*(\underline{\theta}) > x^*(\overline{\theta})$.

When do leaders report truthfully? Competent leaders may mimic incompetent leaders to incur fewer contribution costs, but incompetent leaders may mimic competent leaders to increase their reelection chances. It is thus not clear *ex ante* as to which type's incentive constraint would bind. Instead, a range on Ψ can be found such that both constraints would be simultaneously satisfied:

$$2\gamma \Big(u(x(\overline{\theta}); \underline{\theta}) - u(x(\underline{\theta}); \underline{\theta}) \Big) \leq \Psi \leq 2\gamma \Big(u(x(\overline{\theta}); \overline{\theta}) - u(x(\overline{\theta}), \underline{\theta}) \Big).$$

This range defines the extent to which leaders value electoral benefits relative to utility from exerting effort into public goods. The lower bound on Ψ comes from the competent leader's constraint. By mimicking an incompetent type, a competent leader incurs fewer costs from exerting effort, but misreporting comes at the expense of decreased electoral odds. Conversely, if electoral

incentives are large, incompetent leaders would be incentivized to mimic competent leaders, even if it means exerting more effort than they would prefer to increase the chances of remaining in power.

Evidently, a large swathe of effort recommendations may be incentive compatible, depending on how leaders trade off the value of providing international public goods and maintaining domestic political office. However, since leaders cannot exert infinite amounts of effort into providing public goods, the IO's tools in optimizing this tradeoff are limited. This is particularly difficult when thinking about cases in which leaders care primarily about domestic political survival: if survival concerns become too important, then no incentive compatible mechanism exists.[6]

Lemma 1 *There exists a $\bar{\Psi}$ such that no incentive compatible mechanism exists if $\Psi > \bar{\Psi}$.*

I now winnow the set of incentive compatible recommendations to those that also satisfy obedience constraints: conditional on the public reports of type, leaders must be willing to implement their recommended policies rather than some other profitable deviation. Such a constraint is appropriate as it embodies the idea that no leader ever needs to be a part of the international agreement. The obedience constraint thus requires

$$\underbrace{u(x^*(\theta_i); \theta_i)}_{\substack{\text{policy payoff for} \\ \text{obeying IO as type } \theta_i}} + \underbrace{\frac{\gamma - q + \mu}{2\gamma}\Psi}_{\substack{\text{electoral payoff for} \\ \text{obeying IO as type } \theta_i}} \geq \max_d \underbrace{u(d; \theta_i, \theta_{-i})}_{\substack{\text{policy payoff for} \\ \text{maximal deviation as type } \theta_i}} + \underbrace{\frac{\gamma - q + \mu}{2\gamma}\Psi}_{\substack{\text{electoral payoff for} \\ \text{deviating as type } \theta_i}} \ .$$

The publicity of leader reports nullifies the electoral benefits from exerting costly effort into public goods provision for all leaders *ex post*, regardless of their type. Publicization implies leaders' electoral odds are no longer connected to the effort that they exert into international cooperation. This removes the signaling incentives for leaders to comply with particularly ambitious investments into public goods because the IO washes out the leader's accountability relationship with the voter. The only effort level that satisfies leader i's obedience constraint is her ideal point $\tilde{a}(\theta_i)$. The revelation of information deactivates the electoral mechanism through which leaders exert effort.

Lemma 2 *Recommendations $x^*(\theta_i)$ satisfy obedience constraints if and only if $x^*(\theta_i) = \tilde{a}(\theta_i)$.*

---

[6]If $\Psi > \bar{\Psi}$, leaders always report to the IO that they are competent, $\hat{\theta}_i = \underline{\theta}$. In this case, the IO recommends the largest investment that could be supported by some type of pooling (which is trivially incentive compatible). However, this recommendation is not implementable because it fails the obedience constraint, proved in the appendix.

The IO undermines the accountability chain between leaders and voters, so leaders have no incentives to pledge more ambitiously than their ideal points. Paradoxically, leaders would always pursue greater investments into public goods without the agreement.

Proposition 2 *There exists an equilibrium implemented by the international agreement in which:*

- *leaders truthfully report their types if $\Psi \in [0, \bar{\Psi}]$ and obey recommendations of their ideal points $x^*(\theta_i) = \tilde{a}(\theta_i)$;*

- *competent leaders exert greater effort than incompetent leaders, $x^*(\underline{\theta}) > x^*(\overline{\theta})$;*

- *the value of office-holding has no effect on effort, $\frac{\partial}{\partial \Psi} x^*(\theta_i) = 0$;*

- *expected global effort is less than in the game without the agreement.*

The IO's design, intended to facilitate cooperation via transparency and flexibility, counterintuitively detracts from leaders' domestic incentives to exert effort into public goods provision. By screening for leaders' types and then publicizing this information, the IO makes it so voters no longer condition their retention rule on implemented policy. Compared to the world without the agreement, leaders no longer have electoral motivations to pledge costly contributions.

The appendix also characterizes the conditions under which leaders might want to join such an institution. Intuitively, I find that competent leaders would always prefer this agreement because it facilitates their reelection odds; however, counterintuitively, incompetent leaders only prefer to join if the salience of international cooperation on public goods is low enough that an incompetent leader could win reelection based her valence. International agreements with pledge-and-review frameworks are likely to enjoy broad membership only when cooperation on the relevant international issue is sufficiently unimportant domestically.

Conclusion

This paper zoomed into the effects of two features found in contemporary international cooperative institutions: voluntary commitments and public reporting of information about those commitments. I document a counterintuitive and unintended consequence of institutional design: due to

a multi-layered agency problem, leaders have fewer incentives to exert effort when party to such an international institution. Leaders get caught between signaling competence to their voters, which incentivizes costly investments into global public goods, and screening by international organizations, which generates incentives to understate willingness to contribute. Paradoxically, the transparency facilitated by the agreement demotivates leaders in using effort as a signal of leader type to their voters.

The model also provides an explanation for why unilateral investments into public goods may be rational in the face of global collective action problems, as has been discussed about climate policy (Aklin and Mildenberger 2020; Kennard and Schnakenberg 2023). If domestic publics can learn about leader quality from policy implementation, the opportunity to invest in public goods efforts and signal competence may be valuable for leaders.

References

Aklin, Michaël and Matto Mildenberger. 2020. "Prisoners of the Wrong Dilemma: Why Distributive Conflict, Not Collective Action, Characterizes the Politics of Climate Change." *Global Environmental Politics* 20(4):4–27.

Ashworth, Scott, Ethan Bueno de Mesquita and Amanda Friedenberg. 2017. "Accountability and Information in Elections." *American Economic Journal: Microeconomics* 9(2):95–138.

Dai, Xinyuan. 2002. "Information Systems in Treaty Regimes." *World Politics* 54(4):405–436.

Dai, Xinyuan. 2005. "Why Comply? The Domestic Constituency Mechanism." *International Organization* 59(2):363–398.

Fox, Justin and Stuart V. Jordan. 2011. "Delegation and Accountability." *The Journal of Politics* 73(3):831–844.

Harrison, Rodrigo and Roger Lagunoff. 2017. "Dynamic Mechanism Design for a Global Commons." *International Economic Review* 58(3):751–782.

Harstad, Bård. 2023a. "Pledge and review bargaining." *Journal of Economic Theory* 207(C).

Harstad, Bård. 2023b. "Pledge-and-Review Bargaining: from Kyoto to Paris." *The Economic Journal* 133(651):1181–1216.

Kelley, Judith G. and Beth A. Simmons. 2019. "Introduction: The Power of Global Performance Indicators." *International Organization* 73(3):491–510.

Kennard, Amanda and Keith E. Schnakenberg. 2023. "Comment: Global Climate Policy and Collective Action." *Global Environmental Politics* 23(1):133–144.

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

McAllister, Jordan H. and Keith E. Schnakenberg. 2022. "Designing the Optimal International Climate Agreement with Variability in Commitments." *International Organization* 76(2):469–486.

Prat, Andrea. 2005. "The Wrong Kind of Transparency." *The American Economic Review* 95(3):862–877.

Raiser, Kilian, Başak Çalı and Christian Flachsland. 2022. "Understanding pledge and review: learning from analogies to the Paris Agreement review mechanisms." *Climate Policy* 22(6):711–727.

Rosendorff, B. Peter and Helen V. Milner. 2001. "The Optimal Design of International Trade Institutions: Uncertainty and Escape." *International Organization* 55(4):829–857.

Slechten, Aurélie. 2020. "Environmental Agreements under Asymmetric Information." *Journal of the Association of Environmental and Resource Economists* 7(3):455–481.