The Determinants of Retaliation in International Economic Conflict: A Difference-in-Difference Design

Claas Mertens

University of Oxford

Abstract The primary goal of punitive economic measures (PEMs) such as sanctions is to coerce the target (the sanctioned country) into a policy change. Much is known about the conditions under which coercion is likely to be (un)successful. Although targets frequently adopt punitive economic countermeasures, we know little about what determines a target's decision to retaliate. This article argues that while materially rich senders (sanctioning countries) can rely on their economic might to deter retaliation, relatively weak senders leverage the support of IOs to avoid countermeasures. Furthermore, retaliation is less likely if the target has an inducement to cooperate with the sender. I test the theory with a mixed methods approach. A difference-in-difference analysis with fixed effects and an instrumental variable approach accounting for the potential endogeneity of IO support tests the four arguments quantitatively. Two case studies verify the hypothesized mechanisms in the context of international environmental politics. The findings indicate that the hypothesized determinants have large effects on the likelihood of retaliation in a broad range of contexts and that the hypothesized mechanisms are observable in the two case studies. Especially the finding that IOs disproportionally benefit relatively weak senders has important implications, including for future international environmental policies aiming to achieve cooperation through PEMs.

Key words: economic statecraft · retaliation · international organizations · weak states · international environmental politics

1 Introduction

When do the targets of punitive economic measures (PEMs)[1] retaliate? The use of economic sanctions and other PEMs has steadily increased (Morgan, Bapat, & Kobayashi, 2014). Much of the literature focuses on the conditions under which the sender[2] (the imposing country) can successfully achieve policy concessions from the target (the sanctioned country).[3] These studies distinguish between two outcomes: successful and unsuccessful coercion of the target. The adoption of punitive economic countermeasures by the target, though a relatively frequent response, as this article will show, is largely neglected. This lacuna is unfortunate not only because of its importance for understanding when the initial PEMs are unsuccessful but also due to its interdependence with the occurrence of economic conflict in the first place. On the one hand, retaliation is a source of further economic conflict. The tit-for-tat escalation of the 2018 US-China trade war and the recent Russian restrictions on gas exports to Europe in retaliation for Western sanctions exemplify this. At the same time, the prospect of retaliation can prevent economic conflict by deterring potential senders. Fear of retaliation is, for example, likely to have driven the decision of the European Union (EU) not to adopt economic sanctions against the Russian regime over its intervention in the Syrian conflict in 2016 (Emmott & Guarascio, 2016; Reuters, 2016). Similarly, the EU postponed its planned carbon levy on international flights (an import tariff) in 2012 when 26 countries – including the US, China, and Russia – threatened to retaliate in various ways (Ahmad, 2015; Pauer, 2019).

Despite retaliation's prominent role in policy-making, we know little about when and why targets decide (not) to retaliate. Notable work exists when it comes to trade policy where the goal of the initial measures (usually tariffs) is to gain economically through protectionism. In the context of the WTO, some studies investigate what kind of states retaliate by entering the dispute settlement mechanism in response to the unfavorable trade policy of a foreign country (Bouët & Métivier, 2020; Bown, 2004; Guzman & Simmons, 2005; Mavroidis, Nordström, & Horn, 1999; Sattler, Spilker, & Bernauer, 2014). Similarly, several studies that investigate US trade policy discuss international responses, including countermeasures, but primarily focus on specific US trade legislation (Bayard & Elliott, 1994; Kherallah & Beghin, 1998; Zeng, 2004). However, when it comes to international conflicts where the initial sender's primary goal is not economic gain and protectionism but to coerce the target into making a policy change, only one notable study exists. Peksen and Jeong (2021) link retaliation to the absolute economic wealth of the sender and the target, the level of trade ties between them, the disputed issue, and whether the target is democratic. The study is a helpful starting point but has limitations I aim to address. Theoretically, the authors do not consider the relative power of the sender and target, the role of IOs in avoiding retaliation, and the role of inducements. Empirically, the deployed logit regressions neglect the risk of various biases. I discuss below how this study contributes by addressing these issues.

[1] I refer to 'punitive economic measures' to capture a relatively wide range of measures, including economic sanctions, trade restrictions, and the withdrawal of economic inducements. What distinguishes PEMs from mere trade policy is their goal: while trade policy aims to maximize domestic economic welfare, PEMs aim to influence the policy decision of a foreign government (the target).

[2] I use the terms 'sender' and 'sender coalition' interchangeably as PEMs can be adopted by a single state or a coalition of states.

[3] Drezner (2011) and Aidt, Albornoz, and Hauk (2021) provide good literature reviews on the effectiveness of sanctions.

This paper develops a simple theory that applies to a wide range of scenarios and distinguishes between a target's immediate and future payoffs depending on its decision (not) to retaliate. It focuses on three sets of hypotheses. First, retaliation is more likely if the sender coalition is materially weak compared to the target. Second, the support of an IO for the sender's PEMs reduces the likelihood of retaliation, and this effect is larger for relatively weak senders. As a result, weak senders are more likely to make their adoption of PEMs conditional on IO support. Third, retaliation is less likely if the target has an inducement to cooperate with the sender as opposed to the conflict being zero-sum. Although granting inducements is a helpful tool, especially for weak senders, its use is mostly the privilege of materially rich senders.

I use the Threat and Imposition of Economic Sanctions (TIES) dataset to create a new panel dataset that includes 2,007 episodes of economic coercion with 11 annual observations each (YEAR -5 to YEAR 5), amounting to 22,077 observations. A difference-in-difference (DiD) analysis compares the likelihood of punitive economic (counter)measures by the episode's target against the episode's sender during YEARs in which the senders' initial measures are and are not in place. Fixed effects (FEs) mitigate the risk of various biases. An instrumental variable approach accounts for the potential endogeneity of IO support. The large-n results support the four hypotheses.

The quantitative results support the theory. Most notably, they show that weak sender coalitions disproportionately benefit from IO support and are more likely to obtain it in the first place. This suggests that relatively weak sender coalitions make the adoption of PEMs conditional on IO support as a strategy of avoiding retaliation by powerful targets. I supplement the quantitative analysis with two brief case studies that verify and exemplify the hypothesized mechanisms in the context of international environmental politics. The first one considers the Montreal Protocol, under which a materially powerful coalition of states supported by the United Nations (UN) successfully phased out the global use of Ozone depleting substances through a combination of punitive trade restrictions and inducements. I contrast this with the EU's 2012 adoption of a punitive carbon levy on international flights, which the EU abandoned after a materially powerful coalition of 26 states threatened and adopted countermeasures.

This study makes several contributions to the literature on economic statecraft and beyond. Theoretically, it contributes by focusing on IOs' role in international economic conflict and emphasizing how IOs can strengthen (coalitions of) sender states. It accentuates how materially weak states leverage IO support to take on relatively strong adversaries and avoid retaliation. Furthermore, it shows how different inducements can facilitate cooperation and avoid retaliation. This study makes an empirical and methodological contribution through its innovative deployment of DiD analyses to study retaliation. Most commonly, DiD distinguishes between a treated and a control group but can also be used to distinguish between different treatments (Duflo, 2001; Fricke, 2017). This study, therefore, utilizes DiD to contrast different forms of initial PEMs, like initial measures with and without IO support, constituting a novel approach to studying how and why international conflicts escalate. Finally, the case studies verify the developed theory and apply it to international environmental politics.

Climate change and other environmental challenges are increasingly becoming sources of international economic conflict (Colgan, Green, & Hale, 2021). At the same time, PEMs offer potential solutions to these global challenges (Barrett, 1997; Nordhaus, 2015). This article contributes to the literature on international environmental politics by investigating the conditions under which PEMs are likely to facilitate cooperation vs. retaliation in the context of global emission abatements. The findings have

important implications, for example, for the EU's ongoing preparations for a punitive import tariff on carbon emitted outside its territory, commonly referred to as a border carbon adjustment (BCA).

The next section lays out the argument in detail. Section three introduces the data and the deployed DiD models. Section four presents the results. Section five further tests and exemplifies the theory through two brief case studies.

2 The Determinants of Retaliation

The first of the following subsections discusses the target's payoffs associated with various outcomes depending on its decision (not) to retaliate. The second subsection focuses on the sender's ability to shape these outcomes and thereby avoid retaliation. Together, they explain what determines a target's decision to retaliate.

2.1 Framework: Target's Payoffs of Retaliation

Figure 1 summarizes the target's payoffs depending on its decision (not) to retaliate. The four types of payoffs can be viewed as different currencies that a target can gain or lose when retaliating. Immediate payoffs include both the economic costs and benefits resulting from the measures adopted by each state and payoffs resulting from the disputed policy outcome. For future payoffs, I distinguish between deterrence and legitimacy. A target weighs these four types of payoffs when deciding whether to retaliate.

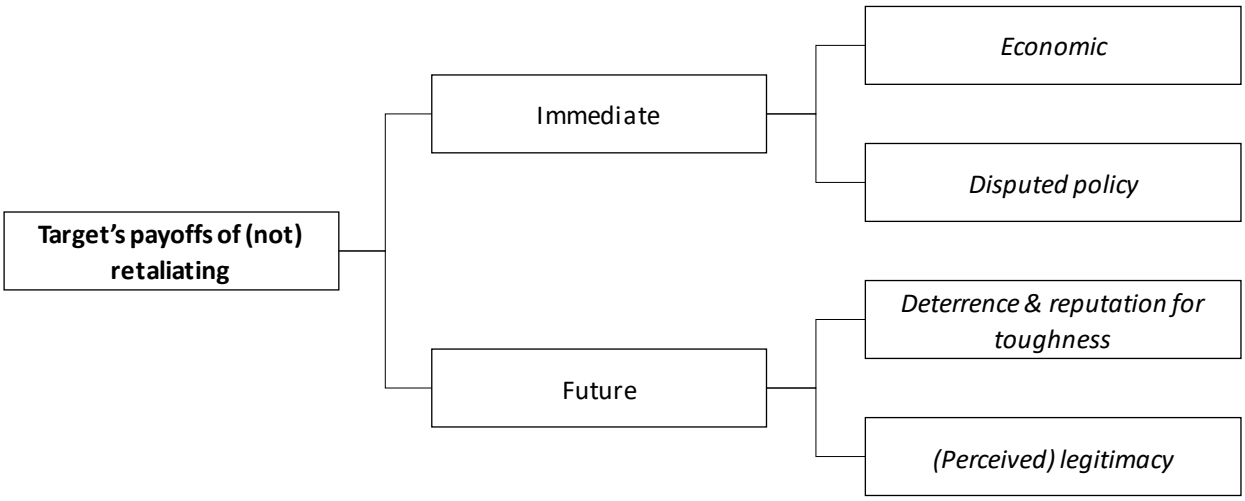


Figure 1. *Payoffs of retaliation framework*

Targeted states derive several immediate benefits from retaliating. Retaliation through protectionist import restrictions benefits certain interest groups within the target economically. For example, in retaliation for Western sanctions in 2014, the Russian regime restricted the import of agricultural produce. Though costly for Russian consumers, the countermeasures' protectionism greatly benefited local business elites (Hedberg, 2018). Furthermore, retaliation through the withdrawal of inducements (like foreign aid) also delivers immediate economic benefits as it saves resources that would otherwise be

transferred to the opposing side. (Unexpected) retaliation may deliver immediate policy benefits by inducing the initial sender to end its measures and retract its demand for a policy change.

Yet, several immediate costs incentivize a target not to retaliate. First, retaliatory measures available to the target may not be economically beneficial but costly instead. While beneficial for some interest groups, import restrictions hurt the consumers/importers of the restricted goods and services. Similarly, other restrictions like financial or export sanctions are costly for both sides, even if costs are asymmetric (Keohane & Nye, 1973). Second, retaliation may provoke additional measures by the initial sender and, in extreme cases, a continuous tit-for-tat escalation with increasing economic costs for both sides. Third, retaliation may lead to less optimal policy outcomes by reducing the likelihood of mutually beneficial cooperation, including collective action.

The initial target may derive future benefits from retaliation by improving its reputation for toughness or deterring the adoption of similar measures in the future. Reputation and deterrence matter, especially when the sender and target are rivals, meaning their expectation of future conflict is high (Drezner, 1999). The Russian retaliation against Western sanctions in 2014 and 2022 and the 2018 US-China trade war exemplify this. A reputation for toughness also matters vis-à-vis domestic audiences, as an escalating conflict with a foreign state may spark a 'rally-round-the-flag' effect (Baker & Oneal, 2001; Mueller, 1970).

Finally, targets may suffer future costs when retaliating. Legitimacy does matter in international relations (Dellmuth, Scholte, Tallberg, & Verhaegen, 2022; Nye, 2004; Stacie, 2018). Depending on the circumstances, various audiences may view retaliation as illegitimate. Audience groups include the international community of states and the retaliating state's domestic constituency.

2.2 The Sender's Ability to Influence the Target's Payoffs

In the following, I apply the above framework to three characteristics of economic conflict that shape the sender's ability to influence the target's payoffs and, thereby, the target's decision (not) to retaliate. My goal is not to be exhaustive but to consider what appear to be the most fundamental characteristics of an economic conflict. I begin with the sender and target states' relative material power. Material power is the classic conception of power in international relations and a helpful starting point (Barnett & Duvall, 2005). I then turn to whether an IO supports the initial punitive measures. This is because, like concepts of power and the role of state actors, IOs play a fundamental role in international relations, including economic conflict. Finally, I turn to different kinds of inducements that the target might have to cooperate with the sender, as these are integral to the target's cost-benefit calculation. The empirical analysis accounts for additional factors like trade ties and enduring rivalries.

2.2.1. Relative Material Power

Escalation is economically costly for both sides. Some interest groups might benefit economically from retaliation, but in the mid to long term, states will only retaliate if they are likely to achieve a beneficial policy outcome or their future payoffs (deterrence and legitimacy) outweigh the increasing economic costs.

States who retaliate tend to do so proportionally to the initial measures. This means that while the two sides suffer comparable absolute economic costs, the relative economic pain of each actor is directly proportional to its relative economic size. In other words, the materially weaker actor suffers

disproportionately (losing $1 is worse if your total wealth is $2 than when it is $20). Of course, this is only a general rule. Independent of the relative economic size, asymmetric economic interdependencies might mean that one actor finds it easier to impose economic costs than the other (Keohane & Nye, 1973). This can, for example, be the case if an actor controls a 'chokepoint' within a network of economic interdependence (Farrell & Newman, 2019). This said, asymmetric interdependencies tend to favor larger economies and are likely only to increase disparities in material power.

H1: Retaliation is more likely if the initial sender is materially weak compared to the target.

Of course, the same logic applies to the initial sender. A relatively weak potential sender is less likely to adopt PEMs in the first place. Despite this bias (which the empirical DiD design will mitigate), relatively weak states do threaten and adopt PEMs against economically stronger opponents. They can have various reasons for this, including reputational considerations vis-à-vis domestic and international audiences. A further reason for relatively weak senders to engage in economic conflict with materially stronger opponents is the support of an IO, which I will discuss next.

2.2.1. IO Support

Institutions play a crucial role in facilitating cooperation in international trade and beyond (Axelrod & Keohane, 1985; Goldstein, Rivers, & Tomz, 2007; Keohane, 2020; Milewicz, 2020). Research on the effectiveness of PEMs has shown that, under certain conditions, the support of an IO increases the likelihood of successful economic coercion (Bapat & Morgan, 2009). IOs play a vital role in coordinating PEMs but typically rely on their member states to implement them, as IOs mostly cannot do so themselves.[4] I distinguish between IOs functioning as forums for member states to coordinate their actions and IO as actors that possess a degree of neutrality. We should expect IO involvement on behalf of the sender to reduce the likelihood of retaliation on both accounts.

As forums, IOs aid coalition-building among states in two ways. First, IOs help coordinate the adoption of PEMs. This coordination increases the sender coalition's efficiency (Abbott & Snidal, 1998) and cohesiveness (Drezner, 2000) and, thereby, its strength. Second, IOs indirectly increase senders' power vis-à-vis the target by signaling the presence of a secondary coalition of potential sender states. Only one or a few member states might adopt the initial punitive measures, but the IO's support signals that a broader coalition of potential sender states may be mobilized if the target escalates the conflict. Suddenly facing this larger coalition of senders increases the target's material costs of escalation considerably.

As actors, IOs possess autonomy that allows them to operate with a degree of neutrality (Abbott & Snidal, 1998). IO support allows sender states to leverage this neutrality. This is because initial senders can claim more credibly that their measures pursue a legitimate (instead of a purely selfish) cause if a seemingly neutral IO supports them. IO support, therefore, gives legitimacy to the senders' initial PEMs and, in turn, makes retaliation by the target appear less legitimate. Thereby, IO support increases the senders' power (through legitimacy) vis-à-vis the target and makes uninvolved states and individuals more likely to side with the sender instead of the target if the conflict escalates. Ultimately, this raises the target's costs of retaliating, making it less likely.

[4] The one big exception to this is the EU which implements most types of PEMs itself, which is why I consider it to be a state actor, not an IO, for the purpose of this study.

H2a: Retaliation is less likely if an IO supports the initial PEMs.

Materially strong and weak senders both gain from IO support. However, materially rich senders are likely to depend less on IO support as they can deter targets from retaliating through their material power. Weaker senders lack this ability, which is why I expect them to benefit disproportionally from IO support:

H2b: If the initial sender is relatively weak, IO support for the initial PEMs is associated with a disproportionately large reduction in the likelihood of retaliation.

For fear of countermeasures, weak senders are generally less likely to adopt PEMs. I expect them to be more likely to make the adoption of PEMs conditional on IO support to avoid retaliation, while strong senders are more likely to adopt PEMs irrespective of IO support.

H2c: PEMs adopted by relatively weak senders are more likely to have IO support.

In sum, materially weak senders are more likely to face retaliation in response to their adoption of PEMs (H1). IO support for the initial measures reduces the likelihood of retaliation for all senders, irrespective of their material capabilities (H2a). However, relatively weak initial senders benefit disproportionally from this effect (H2b). Materially weak senders are, therefore, more likely to make the adoption of initial PEMs conditional on IO support to avoid retaliation (H2c).

2.2.2. Inducements

Material power is not restricted to the ability to hurt. It includes the sender's capacity to influence the target by making the provision of inducements conditional on the target's cooperation. I follow Van Aaken and Simsek (2021) in distinguishing between internal and external inducements that actors may receive when cooperating.

Internal inducements are benefits the sender and the target derive from cooperation on the disputed issue. If PEMs are, for example, aimed at overcoming a prisoner's dilemma or another type of collective action problem, internal inducements are derived from cooperation and make both parties better off. Internal inducements come in addition to the benefits of avoiding PEMs, incentivizing the target to cooperate instead of retaliating. Environmental protection is an issue area where cooperation comes with a relatively high likelihood of internal inducements. Several countries that, for instance, share access to a lake might individually be better off when overfishing or polluting the lake but may at the same time be better off if they all did not. Cooperation and its enforcement through PEMs can therefore deliver internal benefits (like clean waters and sustainable fish populations) to senders and targets alike. However, there are many instances of economic coercion where the target does not receive internal benefits from giving in to the sender's demands. Resolving a territorial dispute, halting a military invasion, accepting regime change, and even improving human rights are often one government's gain but another's loss. Trade disputes, too, are often perceived to be zero-sum.

External inducements can be thought of as side payments linked to the recipient's/target's cooperation. Like internal inducements, they open the bargaining range, making a mutually beneficial outcome more likely and retaliation less likely. Foreign aid and favorable loans offered by one state to another for cooperation are common external inducements. There are two ways in which a sender can leverage external inducements in the context of PEMs. First, the sender can withdraw previously granted or

promised external inducements. Second, the sender may offer an external inducement to the target in case of cooperation and as an alternative to incurring PEMs.

I expect internal or external inducements to reduce the likelihood of retaliation in two ways, irrespective of the sender's and target's relative material power. First, senders can make inducements conditional on the target's cooperation. Second, the presence of inducements is likely to make retaliation appear less legitimate.

H3a: Retaliation is less likely if the target receives internal or external inducements if it cooperates.

As in the case of IO support, I expect relatively weak senders, who, all else equal, face a higher likelihood of retaliation, to disproportionally benefit from linking the target's cooperation to inducements.

H3b: If the initial sender is relatively weak, inducements to cooperate are associated with a disproportionately large reduction in retaliation likelihood.

However, the provision of external inducements requires material resources. The lack of these makes weak senders weak in the first place. As a result, even though weak senders disproportionally benefit from handing out external inducements in terms of avoiding retaliation, they are less likely to do so. Internal inducements, too, are more likely to be associated with materially strong senders, as income is a small but significant predictor of support for issues such as environmental protection (Bergquist, Nilsson, Harring, & Jagers, 2022). Even though beneficial in terms of avoiding retaliation, I expect inducements to be primarily a tool of materially strong senders.

H3c: Relatively weak senders are less likely to provide targets with inducements to cooperate.

In sum, retaliation is less likely if the target has an inducement to cooperate, as opposed to retaliating (H3a). As with IO support, this effect is larger if the sender is relatively weak (H3b). However, unlike in the case of IO support, the provision of inducements is a weapon of the strong, not the weak, making materially weak senders less likely to rely on inducements (H3c).

3 Data & Methods

3.1 The Original Data

To test the hypotheses, I rely on the TIES dataset (Morgan et al., 2014), which captures 1,412 episodes of PEMs between 1945 and 2010. TIES' large number of observations distinguishes it from alternative datasets, which, for example, focus exclusively on a subset of senders like the US, EU, and UN. Neither TIES nor alternative datasets include information on whether the target of the initial PEMs retaliated by adopting punitive economic countermeasures. TIES' extensive coverage is crucial for this study as it enables not only the identification of a wide range of initial episodes but also the identification of episodes that may[5] constitute a retaliatory response by the initial target against the initial sender.

[5] I say "may" as the initial target might have adopted PEMs irrespective of the sender's initial measures in which case these should not be considered retaliatory. I therefore only consider PEMs beyond the base rate likelihood of measures by the target against the sender as retaliatory. I understand the base rate as the likelihood of measures

3.2 New Panel Data & Dependent Variable

In many EPISODES, punitive measures are not adopted by a single sender but by a coalition of sender states. To test for retaliation in response to each of the involved senders, I split these cases into one observation per sender state, leading to a total of 2,007 EPISODES.[6]

I then transform the 2,007 EPISODES into panel data. Figure 2 exemplifies the new data format for EPISODES that started between 2002 and 2004. For each EPISODE, I create five annual observations before the adoption of the initial PEMs (YEAR -5 to YEAR -1) and six yearly observations starting with the adoption of the initial measures (YEAR 0 to YEAR 5). The result is 2,007 sender-target dyads with eleven observations each, amounting to 22,077 (11x2,007) observations.

INITIAL ONGOING is a dummy variable that indicates whether the PEMs by the initial sender against the initial target are ongoing. The variable is coded "0" for all observations from YEAR -5 to YEAR -1. The variable is coded "1" for all observations in YEAR 0, as this is when the 'initial' PEMs are first adopted. For observations from YEAR 1 to YEAR 5, the coding depends on whether the initial measures are ongoing at the beginning of the panel year.[7]

TARGET AGAINST SENDER is the primary dependent variable. This dummy specifies whether the initial target adopts PEMs against the initial sender during any of the initial EPISODE's YEARs (-5 to 5). Naturally, 'retaliation' cannot occur prior to the initial measures in YEAR 0. However, the EPISODE's initial target might have previously adopted punitive measures against the EPISODE's initial sender. If these prior measures were ongoing between YEAR -5 and YEAR -1 but finished before the EPISODE's 'initial' measures started in YEAR 0, the TARGET AGAINST SENDER variable is coded as a "1" (see yellow fields in Figure 2, TARGET AGAINST SENDER = 1 | INITIAL ONGOING = 0).[8] If the sender adopts initial measures (INITIAL ONGOING = 1) and the target adopts its own measures (TARGET AGAINST SENDER = 1) while the sender's initial measures are ongoing, this may[9] constitute a retaliatory response (see red fields in Figure 2).

This coding has the advantage that we can exploit the time variation within EPISODEs and deploy a DiD approach, comparing the likelihood of measures by the TARGET AGAINST the SENDER when the initial sender has put measure against the target in place (INITIAL ONGOING = 1) to times where no initial measures are in place (INITIAL ONGOING = 0). In other words, we can compare the likelihood of measures by the TARGET AGAINST the SENDER within (as opposed to across) the 11-year sender-target dyads. This enables us to estimate the likelihood of measures by the TARGET AGAINST the SENDER, depending on whether and what kind of initial measures were adopted.

---

by the EPISODE's target against the sender during observations where the initial measures are not in place. See *Model Specification* and *Main Results* sections for more details.

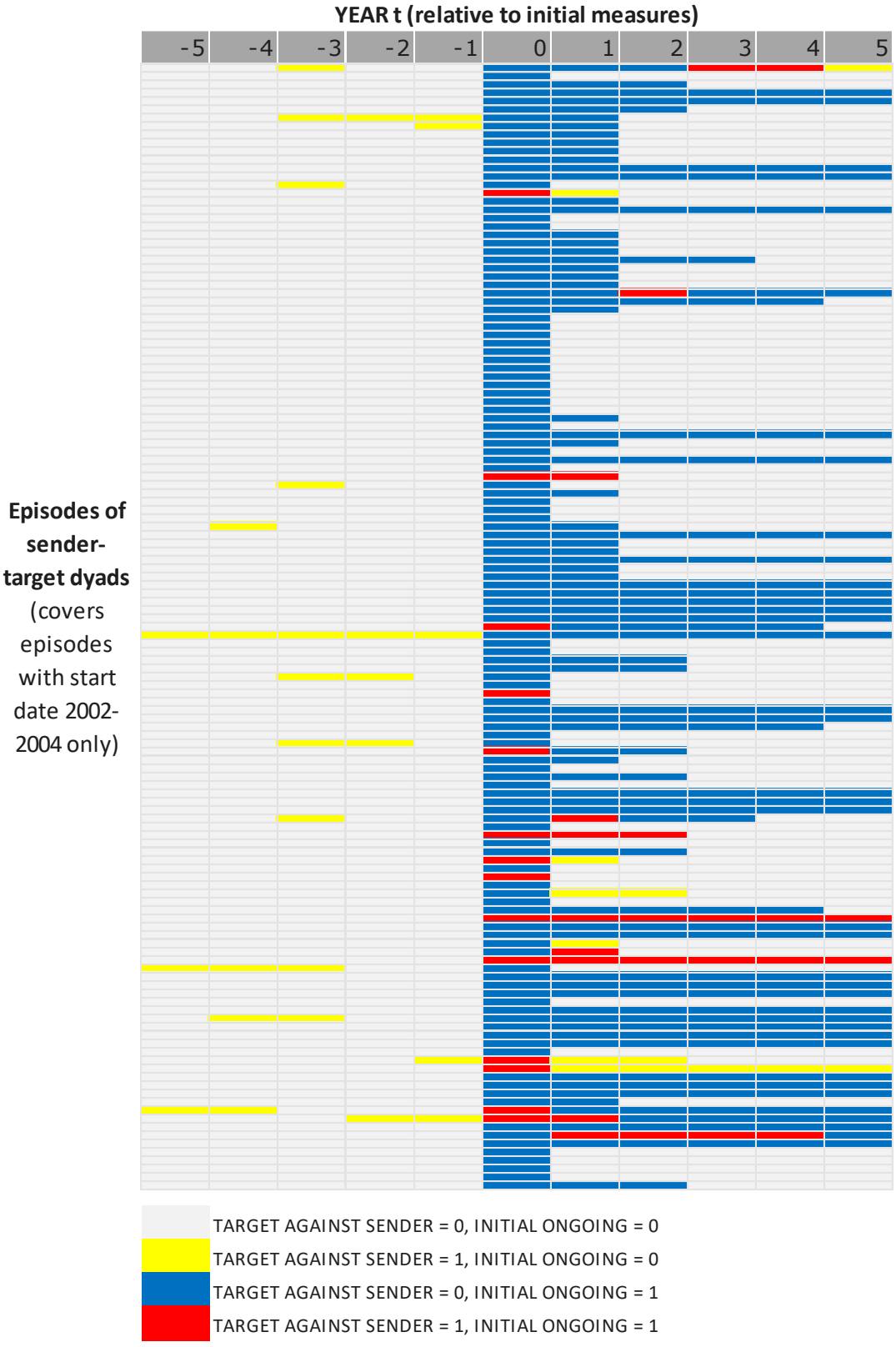
[6] I still run robustness tests with only the original 1,412 episodes.

[7] The dataset includes the threat as well as the imposition of PEMs. For the main analysis I code both the dependent TARGET AGAINST SENDER variable and the INITIAL ONGOING variable as 1 for threat and imposition observations. For both variables I do, however, run robustness tests that count only threat and only imposition observations as measures by the TARGET AGAINST the SENDER and as INITIAL ONGOING.

[8] If the 'prior measures' are still ongoing in YEAR 0, it is the episode's 'initial measures' that are in fact retaliatory. In these cases, the TARGET AGAINST SENDER variable is coded 0, and the retaliation will be picked up in the dataset's episode that covers the 'prior measures'.

[9] See footnote 5.

Figure 2. *Exemplary extract from panel dataset for 2002-2004 and dependent variable coding*



Examples of cases coded as retaliatory are disputes over gas supplies between Russia and Ukraine (2000 -2004), a dispute over the trade of beef between the US and Japan (2003-2005), and the conflict between the US and EU over subsidies to airplane makers that started in the early 2000s.

3.3 Independent Variables

WEAK SENDER is a dummy coded "1" if the aggregated GDP of all sender states involved is smaller than the GDP of the target in YEAR 0. I conduct robustness tests with dummies coded "1" only if the target's GDP is 1.5 and 2 times larger than the sender coalition's. Furthermore, I also conduct a robustness test with a continuous variable obtained by taking the natural logarithm of the GDP ratio of the target and the sender coalition: $Log_e\left(\frac{GDP\ target}{GDP\ sender}\right)$. I rely on GDP data from an updated version of Gleditsch's (2002) dataset (Gleditsch, 2013).

IO SUPPORT is a dummy variable included in the TIES dataset. This variable is coded "1" if an IO was involved on behalf of the sender.[10] The IO need not be the sole sender of the punitive measures. I exclude the dataset's 34 episodes where an IO is the only recorded sender. This is because target states can retaliate against other states but cannot feasibly target an IO. Not excluding the episodes in which an IO is the sole sender would therefore risk biasing the results in favor of H2a, which predicts retaliation to be less likely if an IO supports the initial measures. At the same time, this exclusion means that the impact of IOs on the likelihood of retaliation might not be captured in full. The exclusion does not apply to the EU, which I treat as a state since target states can (and do) retaliate against the EU's punitive measures.

IO X WEAK SENDER is an interaction of the previous two variables. Its purpose is to test whether the effect of IO SUPPORT differs depending on the relative material power of the sender and target (H2b).

INDUCEMENT TO COOPERATE tests the prediction of H3a that targets are more likely to cooperate (and therefore less likely to retaliate) if cooperation comes with internal or external inducement for the target. For the main analysis, I combine the three sources of inducements outlined in the theory: (1) if the initial sender demands improved environmental protection from the target, I consider this to be a case where cooperation tends to offer relatively high internal inducements;[11] (2) if the sender withdraws previously promised or granted inducements, like foreign aid, as a type of PEM, I consider this to be a type external inducement as the previously granted inducement is made conditional on the target's cooperation; (3) if the sender offers an additional inducement to the target if it cooperates, I also consider this as a case of external inducement. I run robustness tests with variables capturing these three scenarios individually.

INDUCEMENT X WEAK SENDER tests H3b's prediction that relatively weak senders disproportionally benefit from handing out inducements in terms of avoiding retaliation.

RIVALRY is a control based on the rivalry dataset by Klein, Goertz, and Diehl (2006). I deploy the dataset's Type 2 rivalry variable, which focuses on enduring political rivalries that typically last for several years

[10] In most cases an IO decided to adopt PEMs against a targeted state. These measures are then adopted and enforced by (some of) the IO's member states. But the bare minimum for this variable to be coded as 1 is that "at least one member of the institution …*[threatens or proposes]*… that the body as a collective adopts sanctions against the target" (Morgan, Bapat, & Kobayashi, 2013).

[11] Disputed issues include wildlife protection, emission reductions, adoption of cleaner technologies, and the control of acid rain (Morgan et al., 2013).

(while Type 1 captures isolated military conflicts that do not require the two actors to be rivals). To determine whether PEMs were adopted during an ongoing rivalry, I convert the rivalry data to a dyadic format and match it with the PEMs' sender, target, and year.

TRADE LINKAGE is a control that approximates the sender's and target's trade dependencies. I follow Peksen and Jeong (2021) in coding it by adding up the target's imports and exports (in current US$) with the sender and dividing the sum by the target's GDP. Import and export data also stems from the Gleditsch (2013) dataset.

3.4 Model Specification

The new panel data format has the advantage of enabling the deployment of a DiD model with FEs. FEs mitigate the risk of time-invariant characteristics of a particular EPISODE biasing the results (including characteristics of the sender, the target, and their relation). This is because EPISODE FEs restrict the model to comparisons between observations within an 11-year sender-target dyad (as opposed to making comparisons across these dyads). In other words, the counterfactual of ongoing initial measures (INITIAL ONGOING = 1) are the observations of the same EPISODE during which initial measures are not ongoing (INITIAL ONGOING = 0). Additionally, I add DECADE FEs to guard against potential biases resulting from systemic changes across time.

Basic DiD models compare a treated and an untreated group. However, the comparison of two different treatments is well established (Duflo, 2001; Fricke, 2017) and well suited to address whether different forms of initial PEMs (i.e., different treatments), such as measures with and without IO SUPPORT or with a relatively WEAK SENDER, affect the likelihood of measures by the TARGET AGAINST the SENDER differently.

In sum, this approach allows us to test in a first instance whether the likelihood of measures by the initial target against the initial sender (TARGET AGAINST SENDER) changes at all with the adoption of initial measures (INITIAL ONGOING), and in a second instance (but within the same model) whether characteristics such as IO SUPPORT or a relatively WEAK SENDER affect the likelihood of measures by the TARGET AGAINST the SENDER.

I use a linear OLS model despite the binary dependent variable. OLS is preferable for models with FEs and interaction effects (Gomila, 2021; Hellevik, 2007), both of which I use. Furthermore, non-linear models are problematic when testing DiD's parallel trends assumption (Lechner, 2011).

Standard errors are clustered by EPISODE as observations within episodes are not random/uncorrelated.

Model 1 starts by testing only the effect of whether initial measures are in place (INITIAL ONGOING) on the dependent variable:

$$TARGET\ AGAINST\ SENDER_{it} = c_1 + \gamma\ INITIAL\ ONGOING_{it} + \sum_{i=1}^{2,007} \delta_i\ EPISODE_i + \varepsilon_{it}$$

In the above equation $TARGET\ AGAINST\ SENDER_{it}$ indicates whether the target adopts PEMs against the initial sender in panel YEAR *t* of EPISODE *i*. Model 2 adds FEs for DECADE *d* and introduces the first explanatory variable of interest, WEAK SENDER (H1). All explanatory variables are interacted with the

INITIAL ONGOING variable to account for treatment only taking place during these observations. The model is specified as follows:

$$TARGET\ AGAINST\ SENDER_{itd} = c_1 + \beta_1\ (\boldsymbol{WEAK\ SENDER}_i \times INITIAL\ ONGOING_{it}) + \gamma\ INITIAL\ ONGOING_{it} + \sum_{i=1}^{2,007} \delta_i\ EPISODE_i + \sum_{d=1940\text{s}}^{2010s} \mu_t\ DECADE_d + \varepsilon_{itd}$$

Models 3-7 are specified similarly to model 2 but vary in terms of included explanatory variables. Model 7 includes all three explanatory variables and the interaction terms:

$$\begin{aligned} TARGET\ AGAINST\ SENDER_{itd} = {} & c_1 + \beta_1\ (\boldsymbol{WEAK\ SENDER}_i \times INITIAL\ ONGOING_{it}) \\ & + \beta_2\ (\boldsymbol{IO\ SUPPORT}_i \times INITIAL\ ONGOING_{it}) \\ & + \beta_3\ (\boldsymbol{IO\ SUPPORT}_i \times \boldsymbol{WEAK\ SENDER}_i \times INITIAL\ ONGOING_{it}) \\ & + \beta_4\ (\boldsymbol{INDUCEMENT\ TO\ COOPERATE}_i \times INITIAL\ ONGOING_{it}) \\ & + \beta_5\ (\boldsymbol{INDUCEMENT}_i \times \boldsymbol{WEAK\ SENDER}_i \times INITIAL\ ONGOING_{it}) \\ & + \gamma\ INITIAL\ ONGOING_{it} + \sum_{i=1}^{2,007} \delta_i\ EPISODE_i + \sum_{d=1940\text{s}}^{2010s} \mu_t\ DECADE_d + \varepsilon_{itd} \end{aligned}$$

3.5 Summary Statistics

Table 1 summarizes the data, differentiating between treated and untreated units (INITIAL EPISODE ONGOING). Furthermore, this study distinguishes between different types of treatment (e.g., initial measures with IO SUPPORT vs. initial measures without IO SUPPORT), which Table 1 reflects.

Table 1. *Summary statistics for treated and untreated observations*

Variable	Obs	Mean	Std. Dev.	Min	Max
Untreated (INITIAL EPISODE ONGOING = 0)					
Dependent variable					
TARGET AGAINST SENDER	16,863	.041	.199	0	1
Independent variables *(variation of treatment)*					
WEAK SENDER	15,344	.228	.42	0	1
IO SUPPORT	16,863	.352	.478	0	1
IO X WEAK SENDER	15,344	.077	.267	0	1
INDUCEMENT TO COOPERATE	16,863	.213	.409	0	1
INDUCEMENT X WEAK SENDER	15,344	.006	.078	0	1
Independent variables *(unrelated to treatment)*					
RIVALRY	16,863	.042	.201	0	1
TRADE LINKAGE	13,573	5.148	8.784	0	89.57
DECADE	16,863	1978.413	16.556	1940	2010
Panel YEAR	16,863	-.449	3.384	-5	5
Treated (INITIAL EPISODE ONGOING = 1)					
Dependent variable					
TARGET AGAINST SENDER	5,214	.072	.259	0	1
Independent variables *(variation of treatment)*					
WEAK SENDER	4,836	.155	.362	0	1
IO SUPPORT	5,214	.398	.49	0	1
IO X WEAK SENDER	4,836	.051	.221	0	1
INDUCEMENT TO COOPERATE	5,214	.269	.444	0	1
INDUCEMENT X WEAK SENDER	4,836	.004	.059	0	1
Independent variables *(unrelated to treatment)*					
RIVALRY	5,214	.065	.247	0	1
TRADE LINKAGE	4,162	5.601	9.808	0	94.37
DECADE	5,214	1981.513	17.037	1940	2010
Panel YEAR	5,214	1.452	1.594	0	5

4 Results

4.1 Parallel Trends Graphs

Figure 3 presents the observed means of the dependent variable (TARGET AGAINST SENDER) across time (YEAR) for each of the three main explanatory variables.[12] Recall that, unlike basic DiD designs, this study does not compare a treated and an untreated group but instead compares different types of treatment. Therefore, the two trend lines in each graph indicate the likelihood of PEMs by the TARGET AGAINST the SENDER and how it changes with the onset of different treatments in YEAR 0 (for example, PEMs by a WEAK SENDER compared to a strong sender).

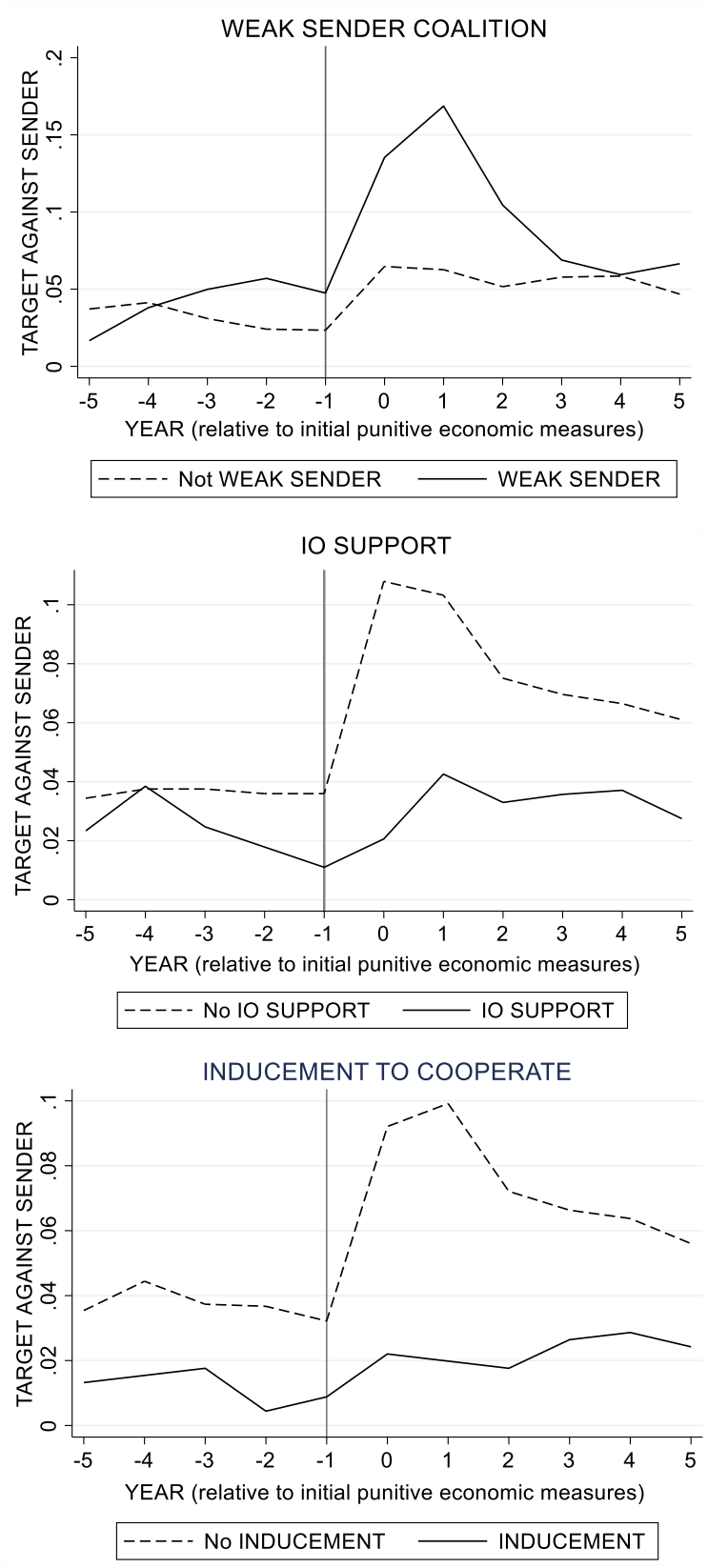
The most important assumption underlying causal inference through DiD designs is that trends prior to treatment are parallel. The three graphs do not indicate that this assumption is not fulfilled. In all three instances, the trends do not appear perfectly parallel in the pretreatment period (YEAR -5 to -1), but this is to be expected when working with observational data and this kind of sample size. Crucially, none of the graphs points towards a systematic divergence of trends prior to treatment.

In episodes involving an INDUCEMENT TO COOPERATE for the target, the pretreatment rate of measures by the TARGET AGAINST the SENDER is consistently below the counterfactual (episodes without an INDUCEMENT TO COOPERATE). Although this points towards a potential pretreatment difference in these two groups, it does not pose a problem for causal inference. This is because pretreatment trends appear parallel and the DiD model, by definition, captures the "differences in the differences," not simple differences, that occur with the onset of treatment.

Besides enabling the inspection of pretreatment trends, the graphs give a first impression of the explanatory variables' average treatment effects (ATE). The first graph, for example, indicates the pretreatment rate of PEMs against a WEAK SENDER to be about 2-5%. This sharply increases to 13-17% in the two years after the onset of treatment (YEARs 0 and 1). This contrasts with the rate of PEMs against relatively strong senders (i.e., not WEAK SENDERs), which the graph shows to increase from about 3-4% prior to the adoption of initial punitive measures to 5-6% in the YEARs posttreatment. The graph, therefore, suggests that both weak and strong senders can expect an increase in PEMs against them in the years after the adoption of initial measures but that this increase is larger for WEAK SENDERs. It is this "difference in difference" that the DiD design captures and that the following section specifies in greater detail.

[12] The interaction terms (IO X WEAK SENDER and INDUCEMENT X WEAK SENDER) cannot be illustrated in this simple way.

Figure 3. *Observed mean rates of measures by TARGET AGAINST SENDER by explanatory variable*



4.2 Main Results

Table 2 presents the main results of the DiD analyses. All coefficients can be easily interpreted as percent change since the dependent variable is binary and the model linear.

The constant of models 1 to 7 indicates that the annual likelihood of measures by the TARGET AGAINST the SENDER in YEARS, during which the INITIAL EPISODE is not ONGOING, is about 4%. This should be understood as the baseline probability of economic conflict. Only economic conflict exceeding this baseline should be considered a retaliatory response to the initial PEMs.

Model 1 includes only the INITIAL EPISODE ONGOING dummy, which indicates whether initial PEMs are in place during a given panel. Together with the model's constant, it forms the counterfactual to which the additional explanatory variables of the subsequent models are compared. The model indicates that if the INITIAL EPISODE is ONGOING, the likelihood of measures by the TARGET AGAINST the SENDER increases by a further 4%, to a total of 8%. Therefore, the model supports the underlying assumption that PEMs by the TARGET AGAINST the SENDER become more likely when initial measures are in place.

Model 2 supports the prediction of H1. Relatively WEAK SENDERs face a combined annual likelihood of countermeasures by the target of 13% (4%+2%+7%), while stronger sender coalitions face (counter)measures with an annual likelihood of only 6% (4%+2%).

Model 3 supports the prediction of H2a. It shows that IO SUPPORT is associated with a 5 percentage point (pp) lower likelihood of countermeasures.

Model 4 includes both previous variables and adds the interaction term IO X WEAK SENDER to test the prediction of H2b that relatively WEAK SENDERs disproportionately benefit from IO SUPPORT. The results support this prediction. The effect size of the WEAK SENDER variable increases from 7% to 11%, indicating that weak senders without IO SUPPORT are even more likely to face retaliation than indicated by model 3. The interaction term (IO X WEAK SENDER) indicates that the support of an IO almost entirely counters the economic weakness of the sender coalition as the likelihood of countermeasures drops by 10 pp. In other words, a weak sender with IO support faces a similar likelihood of retaliation as a strong sender without IO support. With the inclusion of the interaction term, the IO SUPPORT variable's coefficient now indicates the base rate effect of IO support that applies to both relatively strong and weak senders. The size of this effect is reduced by 2 pp (to -3%) but remains statistically significant at the $p<0.05$ level. In sum, the model supports the theory's prediction that both strong and weak senders benefit from IO SUPPORT, but that weak senders benefit more.

Model 5 supports the prediction of H3a that the presence of internal or external inducements (INDUCEMENT TO COOPERATE) is associated with a lower likelihood of countermeasures. The effect size of -5% is similar to the 5 pp increase associated with the INITIAL EPISODE being ONGOING, indicating that the adoption of initial punitive measures that come with internal or external inducements for the target is associated with almost no increase in the likelihood of measures by the TARGET AGAINST the SENDER compared to no initial measures being in place.

Model 6 adds the interaction term INDUCEMENT X WEAK SENDER and supports the prediction of H3b that relatively weak senders disproportionally benefit in terms of avoiding retaliation if the target has an inducement to cooperate.

Table 2. *DiD main results; dependent variable: measures by TARGET AGAINST SENDER*

VARIABLES	(1) *BASIC*	(2) *WEAK SENDER*	(3) *IO*	(4) *IO INTERACTION*	(5) *INDUCE. TO COOPERATE*	(6) *INDUCE. INTERACTION*	(7) *FULL*	(8) *INSTRUMEN. VARIABLE IO*
WEAK SENDER (H1)		**0.072*** **		**0.113*** **		**0.068*** **	**0.111*** **	**0.122*** **
		(0.017)		**(0.025)**		**(0.018)**	**(0.026)**	**(0.022)**
IO SUPPORT (H2a)			**-0.046*** **	**-0.028***			**-0.029***	**-0.031***
			(0.010)	**(0.011)**			**(0.011)**	**(0.013)**
IO X WEAK SENDER (H2b)				**-0.103*** **			**-0.106*** **	**-0.135** **
				(0.030)			**(0.030)**	**(0.050)**
INDUCEMENT TO COOPERATE (H3a)					**-0.046*** **	**-0.028***	**-0.029***	**-0.029*** **
					(0.010)	**(0.011)**	**(0.011)**	**(0.009)**
INDUCEMENT X WEAK SENDER (H3b)						**-0.137*** **	**-0.165*** **	**-0.171** **
						(0.031)	**(0.046)**	**(0.054)**
INITIAL EPISODE ONGOING	0.041***	0.023***	0.053***	0.032***	0.047***	0.031***	0.041***	0.042***
	(0.006)	(0.006)	(0.008)	(0.009)	(0.007)	(0.008)	(0.010)	(0.007)
EPISODE FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DECADE FEs	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Instrum. var. (IO SUPPORT and IO X WEAK SENDER)	No	No	No	No	No	No	No	Yes
Constant	0.041***	0.040***	0.041***	0.042***	0.039***	0.040***	0.042***	^
	(0.001)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	
Observations	20,922	20,350	20,922	20,350	20,922	20,350	20,350	20,350
Number of EPISODES	1,902	1,850	1,902	1,850	1,902	1,850	1,850	1,850

Robust standard errors in parentheses; standard errors clustered by EPISODE; *** p<0.001, ** p<0.01, * p<0.05; ^no constant is reported because the two-stage instrumental variable model means it cannot be calculated accurately

Model 7 includes all the previously discussed variables. This has little impact on the size and statistical significance of the previously described effects.

Model 8 adds two instrumental variable estimations to predict IO SUPPORT and IO X WEAK SENDER. This accounts for the possibility that, instead of reducing the likelihood of retaliation through their support, IOs are more likely to support PEMs that are less likely to spark retaliation (regardless of IO SUPPORT). I use the sum of the senders' IO memberships (IO MEMBERSHIP COUNT) as an instrument to predict IO SUPPORT in the model's first stage.[13] The exclusion restriction is the most important condition for the validity of an instrumental variable estimation. An instrument (in this case IO MEMBERSHIP COUNT) is valid if it predicts the instrumented variable (i.e., the dependent variable of stage 1, in this case, IO SUPPORT) but is uncorrelated with the dependent variable of stage 2 (in this case TARGET AGAINST SENDER). Besides being theoretically plausible, the fulfillment of the exclusion restriction is supported empirically. Appendix A reports the results of the stage 1 models and shows that the positive correlation between IO SUPPORT and IO MEMBERSHIP COUNT is statistically highly significant (|p|=0.000). To test that the IO MEMBERSHIP COUNT variable is not correlated with the dependent variable of stage 2 (TARGET AGAINST SENDER), I add it to model 7 (i.e., the full model without an instrumental variable). The results are reported in Appendix B and show no significant correlation between IO MEMBERSHIP COUNT and TARGET AGAINST SENDER (|p|=0.738), supporting the instrument's validity.

The stage 2 results of model 8 support the robustness of models 1-7 as the size and statistical significance of the effects barely change. The two-stage approach of model 8 has a further advantage. The stage 1 model that predicts IO SUPPORT (see Appendix A) indicates that WEAK SENDERs are significantly more likely to have IO SUPPORT than their more powerful peers (H2c). Importantly, this points to IO SUPPORT being a strategy that the weak actively pursue when taking on relatively powerful targets. While economically rich senders can rely on their material power to deter retaliation, weak senders cannot and therefore make the adoption of PEMs conditional on IO SUPPORT. The results of stage 2 indicate that this strategy is highly effective.

To test H3c's prediction that, though particularly beneficial for weak senders (H3b), inducements are predominantly tools of materially strong senders, I run a series of logit regressions (see Appendix D). Across models, the results support the hypothesis, showing that initial PEMs adopted by materially weak senders are less likely to come with inducements for the initial target.

4.3 Robustness Tests

To further test the robustness of model 8, I control for the type of disputed issue. TIES classifies episodes into 15 categories of disputed issues.[14] The results show that both the likelihood of IO SUPPORT (stage 1) and the likelihood of countermeasures by the TARGET AGAINST the SENDER (stage 2) vary across issue

[13] I add an interaction of the IO MEMBERSHIP COUNT and the WEAK SENDER variable as an instrument to account for the interaction term (IO X WEAK SENDER) of the stage 2 model. The result are two stage 1 models, one to predict IO SUPPORT and one predicting IO X WEAK SENDER.

[14] The categories are: Contain Political influence, Contain Military Behavior, Destabilize Regime, Release Citizens, Property, or Material, Solve Territorial Dispute, Deny Strategic Materials, Retaliate for Alliance or Alignment Choice, Improve Human Rights, End Weapons/Materials Proliferation, Terminate Support of Non-State Actors, Deter or Punish Drug Trafficking Practices, Improve Environmental Policies, Trade Practices, Implement Economic Reform, and Other (Morgan et al., 2013).

categories. However, including these additional controls has little effect on the main results. The only notable difference compared to model 8 is that IO SUPPORT is now only significant at the |p|<0.1 level, while IO X WEAK SENDER is now significant at the |p|<0.01 level (see Appendix C).

Various additional robustness tests that build on model 7 further support the results. The following three robustness tests are reported in Appendix E. First, the results are robust when limiting the data to the TIES dataset's original 1,412 episodes instead of splitting episodes involving several sender states into separate observations (see *The Original Data* section). Second, the results are robust when limiting the dependent variable (TARGET AGAINST SENDER) to episodes and observations where PEMs were only threatened or only imposed (instead of combining observations of threats and impositions). Third, including the controls RIVALRY and TRADE LINKAGE does not substantially change the results.

Fourth, the results are also robust when the definition of what constitutes a WEAK SENDER is changed (see Appendix F for full results; see the *Independent Variables* section for the various alternative codings of the variable). Fifth, I break down the INDUCEMENT TO COOPERATE variable into its three components (see Appendix G for full results). The two dummies that capture only environmental issues (internal inducement) and only foreign aid suspensions (external inducement) remain statistically significant. The effect size of the dummy indicating whether the sender promised an additional external inducement if the target cooperates does not change much but misses conventional levels of statistical significance. Finally, I run models 1 to 7 using a logit instead of an OLS model (Appendix H). Although OLS is the more appropriate choice for several reasons (see *Model Specification* section), this is to account for the argument that a logit model better fits the binary dependent variable. All results support the robustness of the main results.

5 A Tale of Two Cases: Economic Coercion and Retaliation in International Environmental Politics

This section exemplifies the developed theory and complements the quantitative analysis by verifying the theory's underlying causal mechanisms. Furthermore, this section applies the developed theory to international environmental politics. Global challenges ranging from climate change to the pollution and overfishing of the oceans are becoming increasingly salient. International environmental challenges are typically conceptualized as collective action problems with (the threat of) PEMs offering a solution to achieving and sustaining cooperation (Barrett, 1997; Nordhaus, 2015). At the same time, PEMs over environmental issues result from an increasing battle between the owners of 'climate-vulnerable' assets like seaside property and 'climate-forcing' assets like fossil fuel production (Colgan et al., 2021). Both perspectives predict that environmental politics will increasingly become a source of international economic conflict.

I start by considering the Montreal Protocol on Substances that Deplete the Ozone Layer, which did not spark a retaliatory response and is considered one of the most successful international treaties (Barrett, 2007; Green, 2009). It relies on a combination of PEMs and inducements (H3) to sustain cooperation, had a relatively strong coalition of initial signatory/sender states (H1), and enjoys UN support (H2). I contrast this case with the EU's attempt to introduce a carbon tariff on flights to and from its territory in 2012, which it was forced to postpone due to severe (threats of) retaliation. The EU was materially weak

compared to the coalition of target states that retaliated (H1). Furthermore, the EU lacked IO support (H2), and I argue that the EU itself should not be considered an IO in the context of PEMs. Finally, the EU exclusively relied on PEMs, instead of a combination of PEMs and inducements (H3).

5.1 The Montreal Protocol on Substances that Deplete the Ozone Layer

"*Perhaps the single most successful international environmental agreement to date has been the Montreal Protocol*" - Former UN Secretary-General, Kofi Annan

Emissions of chlorofluorocarbons (CFCs) were found to harm the earth's Ozone layer in the 1970s. A hole in the Ozone layer was first discovered above Antarctica in 1985, further increasing urgency. CFCs had many advantages over previously used chemicals due to their low toxicity and flammability and were widely used, for example, in refrigerators, air conditions, and deodorant and spray paint aerosol cans. However, when linked to increasing holes in the Ozone layer, which in turn were linked to various adverse health outcomes like skin cancer and immune suppression, the need to drastically reduce CFCs' use motivated the Montreal Protocol.

The developed countries that were early proponents of CFC abatements stood to gain disproportionately – in absolute terms due to their primarily white population for which the risk of skin cancer was particularly high and in relative terms due to their higher willingness to pay resulting from higher GDP per capita (Barrett, 2007, p. 80). It is, therefore, no coincidence that it was Belgium, Canada, Norway, Sweden, and the United States who first took unilateral action in the 1970s to reduce their CFC usage. However, it was also clear that abatements in these countries alone would not be sufficient, among other things, because there was 'leakage' due to production shifting to unrestricted regions. The need for a multilateral global approach motivated the 1987 Montreal Protocol. Initially, the early participants attempted to incentivize participation through PEMs that restricted the trade of CFCs with non-signatories. Although they did not spark retaliation, the PEMs were not sufficient to incentivize many countries to join the Montreal Protocol and phase out the use of CFCs. India initially increased its production of CFCs to take advantage of the reduced supply to many other developing countries due to the Montreal Protocol (Barrett, 2007, p. 80). China and many other developing countries did not join the initial agreement either.

Things changed when developed countries pledged to pay developing countries' incremental cost of phasing out CFCs as part of an amendment of the Montreal Protocol in London in 1990. As a result of this external inducement (H3), developing countries were no longer economically worse off when joining the agreement and additionally got to enjoy the internal inducements from increased Ozone protection. The Multilateral Fund, established for this purpose, received contributions of $4.4 billion between 1991 and 2021 and currently considers 147 out of the 197 countries that are party to the Montreal Protocol as potential beneficiaries (Multilateral Fund, 2022). The Montreal Protocol's success is primarily attributed to these external inducements, which convinced many developing countries, including China and India, to join the agreement from the early 1990s onwards. Developed countries, too, were better off as the internal benefits of increased Ozone protection far exceeded the costs of providing external inducements to developing countries.

As the theory developed in this article would predict, the trade restrictions adopted by the early signatories of the Montreal Protocol did not provoke any retaliatory threats or actions from non-participants, even before external inducements were offered to developing countries. The initial coalition

of states that agreed to phase out CFCs in 1987 and that adopted PEMs against non-participants was by no means weak in terms of material power (H1). The US, the EU, Russia, Japan, Canada, and Mexico were just the largest of the original 46 signatories (United Nations, 2022). Together, these five states and the EU accounted for an aggregated 75% of world GDP in 1987, whereas India and China accounted for approximately 1.6% each (World Bank, 2022).

The materially strong sender coalition of early signatories enjoyed the support of the UN, which served as a forum for building the initial coalition and negotiating later amendments to the Montreal Protocol (H2a). H2b posits that relatively weak senders disproportionally benefit from an IO's support by avoiding retaliation. The Montreal Protocol is the only UN treaty ratified by all 198 UN member states (UN Environment Programme, 2022), and the UN undeniably played a fundamental role in facilitating the Montreal Protocol's steady process through regular amendments over many years. It is likely to have added legitimacy and credibility to the process, which contributed to the Montreal Protocol's success. As always, we cannot observe the counterfactual. It does seem plausible, though, that due to the initial signatories' enormous material power vis-à-vis its targets and despite the UN's important role in facilitating cooperation, the initial signatories did not require UN support to deter retaliation (H2b).

Internal and external inducements to cooperate played a crucial role in the Montreal Protocol's success. The case illustrates how inducements and cooperation were an iterative process in which internal inducements brought together an initial coalition, which then deployed a combination of internal and external inducements and PEMs to increase participation further, leading to a situation in which non-participants had no incentive to retaliate even if they could have done so effectively (H3a). However, the case also illustrates how the provision of external inducements, which played a key role in the Montreal Protocol's success, was made possible only by the initial sender/signatory coalition's material resources that far exceeded the resources of the initial targets/non-signatories (H3c).

5.2 The EU Aviation Directive

In 2008, the EU announced its Aviation Directive. Besides flights within its territory, the directive meant that the EU would subject cross-border flights to carbon pricing through its cap-and-trade Emissions Trading System (ETS) from 2012 onwards. The levy was non-discriminatory as it equally priced the emissions of EU and foreign airlines. Nevertheless, it fulfilled the criteria of a PEM as it attempted to influence the policy of foreign states. The directive not just covered distances flown within the EU's territory but also the part of the journey beyond the EU's border. Therefore, the EU's directive attempted to influence the policymaking of foreign states by incentivizing them to limit carbon emissions from aviation. The EU clarified that it would lift its PEM if a foreign state adopted its own regulation with "an environmental effect at least equivalent" to the carbon emissions pricing through the EU's Aviation Directive (European Parliament and Council, 2008, Art. 17).

International reactions to the implementation of the EU Aviation Directive in 2012 were very different from those of the Montreal protocol. The so-called 'coalition of the unwilling' – consisting of 26 countries, including the US, China, Canada, Russia, and India – emerged and threatened to retaliate in various ways (Ahmad, 2015; Pauer, 2019). These states jointly issued the Moscow Declaration, which included threats to retaliate by (a) "imposing additional levies/charges on EU carriers/ aircraft operators as a form of countermeasure," (b) restricting European airlines' activity in the retaliating states, (c) instructing their airlines not to comply with the EU's directive, and (d) vaguer references towards the risk of an escalating

trade war (GreenAir, 2012; Russian Aviation, 2012). Ahmad (2015) provides a long list of additional threatened and imposed countermeasures by individual states. These include the Chinese suspension of the purchase of 45 Airbus planes with a total worth of $14 billion (Hepher, 2012), Russian refusal to grant additional overflight rights to EU airlines (Reuters, 2012), Canadian threats to place limitations on polar flights by EU carriers (Ahmad, 2015), and US President Obama's signing of a bill giving the US Secretary of Transport the right to prohibit US carriers from complying with the EU's directive (112th Congress, 2012). In response to the overwhelming international countermeasures, the EU announced the suspension of its Aviation Directive in 2013 (European Parliament and Council, 2013) and later limited the directive's scope to flights within the European Economic Area (European Parliament and Council, 2014).

It is impossible to say whether the EU's deployment of a PEM in pursuit of international cooperation would have triggered the same retaliatory response under different circumstances. However, the case is in line with the outlined theory and the quantitative findings of this article. Though generally no economic dwarf, the EU was a weak sender relative to the 26 states that threatened retaliation (H1). The EU's 2012 GDP ($14.6 trillion) amounted to less than half of the aggregated $30.7 trillion GDP of just the five largest targets (US, China, Canada, Russia, and India) (World Bank, 2022). As a result, the EU lacked sufficient material power to deter its targets from retaliating.

This paragraph argues that, in the realm of foreign economic policy, the EU should be considered a state actor, not an IO, and that the case, therefore, fits the prediction of H2a and H2b that relatively weak senders (as the EU was) disproportionally benefit from IO support (which the EU lacked). When it comes to military or fiscal policy, for example, the EU has the character of an IO, as member states may or may not coordinate their actions through the EU. However, the EU has the characteristics of a state actor, not an IO, in the realm of foreign economic policy. The EU Commission, the EU Council, and the EU Parliament act on behalf of the EU's member states, drafting and imposing economic sanctions and other PEMs like the Aviation Directive. Due to the EU Council's unanimity requirement, PEMs always have the backing of all member states. In the theory section, I argued that IO support (which the EU lacked in this case) strengthens the sender coalition and reduces the likelihood of retaliation for two reasons. First, IO support offers a secondary coalition of potential senders that are IO members but do not implement the initial PEMs. This does not apply to the EU, as the initial punitive measures already have the backing of all member states. Second, I argue that IOs' partial neutrality increases the legitimacy of the initial PEMs, thereby making retaliation (appear) less legitimate and further reducing its likelihood. The EU does not benefit from this legitimacy effect as it acts in the interest of all its members when adopting PEMs and therefore lacks IOs' typical neutrality. In line with the theory, the EU's Aviation Directive should therefore be seen as lacking the support of an IO (H2a).

We cannot be sure how things would have played out if the EU had enjoyed IO support. However, it seems plausible that the EU would have benefited disproportionally from IO support precisely because it was a relatively weak sender that lacked the material power to deter retaliation. By acting through a supportive IO, the EU may have had a chance to broaden its (secondary) sender coalition and increase the perceived legitimacy of its PEMs (H2b).

H3a and H3b posit that retaliation is less likely if the target has an internal or external inducement to cooperate with the sender and that this effect is particularly pronounced for relatively weak senders (which the EU was). The EU's Aviation Directive relied on PEMs to achieve cooperation but did not offer any external inducements comparable to those that boosted cooperation under the Montreal Protocol.

Internal inducements to cooperate did exist, as the Aviation Directive's goal was to limit climate change by reducing carbon emissions. But the internal inducement stemming from curbing climate change may have been too small, distant, and uncertain to prevent retaliation. Back in 2012, climate change was less salient than it is today. Furthermore, the cost-benefit ratio of reducing carbon emissions is less extreme in favor of abatements than in the case of Ozone depleting substances (Barrett, 2007, pp. 74-102). Therefore, the targets' overall inducements to cooperate with the EU were relatively low, making retaliation more likely (H3a), especially as the EU was relatively weak (H3b). Finally, precisely because the EU was relatively weak in terms of material resources, it could not feasibly incentivize its targets (who had more than double its GDP) through external inducements (H3c).

5.3 Summary and Policy Implications

The EU's Aviation Directive and the Montreal Protocol shared the intention to achieve cooperation on global emission reductions through the deployment of PEMs. However, outcomes could hardly have been more different. While the Montreal Protocol continues to be celebrated as one of the most successful cases of international cooperation 35 years after its birth, the EU was forced to abandon its international carbon levy in response to the threat and adoption of severe retaliatory measures. As always, we can only speculate on 'what would have happened if,' but the two cases fit this paper's theory. While the EU was a relatively weak sender compared to its targets, the initial senders/signatories of the Montreal Protocol were not (H1). While the EU did not have the backing of an IO, the Montreal Protocol was facilitated by the UN (H2a). While it seems plausible that the EU would have benefited disproportionately from the support of an IO due to its relative economic weakness, retaliation against the Montreal Protocol's early senders/signatories appears unlikely even in the absence of IO support (H2b). Finally, while the EU may have struggled to offer sufficient external inducements to its targets and internal inducements might have been perceived to be minor, the Montreal Protocol's success stems from an effective combination of PEMs and external inducements as well as high internal inducements (H3a-c).

The quantitative and qualitative findings have important implications for future international environmental policymaking. Carbon pricing is crucial for efficiently reducing emissions (Hepburn, Stern, & Stiglitz, 2020; Nordhaus, 2015). One global carbon price would be the economically most efficient way to achieve global abatements but, for now, appears politically unfeasible. If carbon prices vary significantly across regions, carbon 'leakage' will likely lead to competitive disadvantages for firms subjected to higher carbon prices while global emissions are not meaningfully reduced. Border carbon adjustments (BCAs) are a potential solution. BCAs are import tariffs that are levied to account for differences in carbon pricing between regions. BCAs have two main goals: mitigating competitive disadvantages and inducing foreign actors to adopt carbon pricing. Influential studies of the effectiveness of BCAs assume that retaliation can be 'prohibited' (Helm, Hepburn, & Ruta, 2012; Nordhaus, 2015). Though justifiable, this assumption is often unrealistic, as this study has shown. Focusing on the role of trade relations, Böhringer, Carbone, and Rutherford (2016) investigate the conditions under which it would be economically rational for states to retaliate in response to BCAs but neglect political factors. This article makes an important contribution to the study of BCAs by drawing on past cases of economic conflict and highlighting the importance of political factors – first and foremost, IO support, especially for relatively weak senders – for understanding when retaliation is likely.

In July 2021, the EU announced the adoption of a BCA on imports of cement, iron and steel, aluminum, fertilizers, and electricity (European Commission, 2021), which is set to start pricing carbon imports in 2026 (European Parliament, 2022). The EU's renewed attempt to price foreign carbon emissions may be a steppingstone towards the global reduction of carbon emissions. However, this article's findings also provide grounds for caution: the EU would reduce the likelihood of retaliation and increase the chances of success by finding materially powerful partners to join its initial sender coalition, striving for the support of IOs, and providing external inducements to developing countries to offset their costs of complying with the BCA.

6 Conclusion

Economic coercion is playing an increasingly important role in international politics. Much is known about the conditions under which PEMs are likely to achieve policy concessions from their target. However, coercion attempts can be unsuccessful in different ways, and these are understood much less. While a target may choose not to acquiesce and live with the adopted PEMs, retaliation through punitive economic countermeasures is an alternative and relatively frequent response. Until now, we knew little about what determines a target's decision (not) to retaliate. This is unfortunate for several reasons. Retaliation can be a source of economic conflict. At the same time, the risk of retaliation can also be an incentive to refrain from adopting PEMs, especially for materially weak potential senders. Furthermore, the world is increasingly facing global challenges like climate change, overfishing, and pollution of the oceans that can only be addressed through collective action. It is difficult to imagine global solutions to these challenges that do not involve (the threat of) PEMs against free riders. However, PEMs risk backfiring if they provoke retaliation instead of inducing cooperation.

This article has presented a novel statistical approach to studying the determinants of retaliation, mitigating the risk of various biases through the deployment of a DiD design, FEs, and instrumental variable techniques. The model tested hypotheses derived from what appeared to be the theoretically most appropriate starting points for investigating the determinants of retaliation (relative material power, IO support, and inducements to cooperate). There are many more characteristics of economic conflict that this article has left unexplored. While prior research, for example, suggests that trade dependencies increase the likelihood of retaliation (Peksen & Jeong, 2021), initial results obtained through the developed DiD model suggest that, while trade linkage is a predictor of the base rate likelihood of economic conflict, it does not impact the likelihood of retaliation beyond this base rate. But further theoretical and empirical work on this and other potential determinants of retaliation is needed. The developed statistical approach and panel data will help with this.

This article finds that materially strong senders can rely on their economic might to deter retaliation, while relatively weak senders are more likely to face punitive economic countermeasures. However, IO support appears to mitigate materially weak senders' increased likelihood of facing retaliation, pointing to two underlying mechanisms. As forums, IOs give power to senders by enabling stronger coalition-building. As actors, IOs can strengthen senders vis-à-vis the target by giving legitimacy to the adopted measures.

Studies of economic statecraft mostly take the perspective of the US and the EU, which both have a GDP far exceeding that of most targets. But the rise of China and other developing countries has gradually led

to a more multipolar world. This trend is set to continue, which means that even the formerly dominant US and EU will increasingly become relatively weak senders and face retaliation when acting unilaterally. This has important implications for economic statecraft in general but also for international environmental politics. PEMs will likely play an essential role in addressing global challenges. Senders striving to incentivize cooperation through PEMs would do well to (1) find like-minded partners to join their initial coalition, (2) involve IOs whenever possible to strengthen their coalition and increase legitimacy, and (3) offer external inducements to materially weaker states in exchange for their cooperation.

References

112th Congress. (2012). European Union Emissions Trading Scheme Prohibition Act of 2011. Retrieved from https://www.congress.gov/bill/112th-congress/senate-bill/1956/text

Abbott, K. W., & Snidal, D. (1998). Why States Act through Formal International Organizations. *The Journal of conflict resolution, 42*(1), 3-32. doi:10.1177/0022002798042001001

Ahmad, M. T. (2015). Evaluating the effectiveness of the European Union emissions trading system to reduce emissions from international civil aviation. *The McGill international journal of sustainable development law and policy, 11*(1), 115.

Aidt, T. S., Albornoz, F., & Hauk, E. (2021). Foreign Influence and Domestic Policy. *Journal of economic literature, 59*(2), 426-487. doi:10.1257/jel.20201481

Axelrod, R., & Keohane, R. O. (1985). Achieving Cooperation under Anarchy: Strategies and Institutions. *World Pol, 38*(1), 226-254. doi:10.2307/2010357

Baker, W. D., & Oneal, J. R. (2001). Patriotism or Opinion Leadership?: The Nature and Origins of the "Rally 'Round the Flag" Effect. *The Journal of conflict resolution, 45*(5), 661-687. doi:10.1177/0022002701045005006

Bapat, N. A., & Morgan, T. C. (2009). Multilateral Versus Unilateral Sanctions Reconsidered: A Test Using New Data. *International Studies Quarterly, 53*(4), 1075-1094. doi:10.1111/j.1468-2478.2009.00569.x

Barnett, M., & Duvall, R. (2005). *Power in global governance*. Cambridge, UK: Cambridge University Press.

Barrett, S. (1997). The strategy of trade sanctions in international environmental agreements. *Resource and energy economics, 19*(4), 345-361. doi:10.1016/S0928-7655(97)00016-X

Barrett, S. (2007). *Why Cooperate?: The Incentive to Supply Global Public Goods* (Vol. 9780199211890): Oxford University Press.

Bayard, T. O., & Elliott, K. A. (1994). *Reciprocity and retaliation in U.S. trade policy*. Washington, D.C: Institute for International Economics.

Bergquist, M., Nilsson, A., Harring, N., & Jagers, S. C. (2022). Meta-analyses of fifteen determinants of public opinion about climate change taxes and laws. *Nature climate change, 12*(3), 235-240. doi:10.1038/s41558-022-01297-6

Böhringer, C., Carbone, J., & Rutherford, T. (2016). The Strategic Value of Carbon Tariffs. *American Economic Journal. Economic Policy, 8*(1), 28-51. doi:10.1257/pol.20130327

Bouët, A., & Métivier, J. (2020). Is the dispute settlement system, "jewel in the WTO's crown", beyond reach of developing countries? *Review of World Economics, 156*(1), 1-38. doi:http://dx.doi.org/10.1007/s10290-019-00359-w

Bown, C. P. (2004). Trade disputes and the implementation of protection under the GATT: an empirical assessment. *Journal of International Economics, 62*(2), 263-294. doi:https://doi.org/10.1016/S0022-1996(03)00053-9

Colgan, J. D., Green, J. F., & Hale, T. N. (2021). Asset Revaluation and the Existential Politics of Climate Change. *International Organization, 75*(2), 586-610. doi:10.1017/S0020818320000296

Dellmuth, L., Scholte, J. A., Tallberg, J., & Verhaegen, S. (2022). *Citizens, elites, and the legitimacy of global governance*: Oxford University Press.

Drezner, D. W. (1999). *The sanctions paradox: Economic statecraft and international relations*. Cambridge: Cambridge University Press.

Drezner, D. W. (2000). Bargaining, Enforcement, and Multilateral Sanctions: When Is Cooperation Counterproductive? *Int Org, 54*(1), 73-102. doi:10.1162/002081800551127

Drezner, D. W. (2011). Sanctions Sometimes Smart: Targeted Sanctions in Theory and Practice. *International Studies Review, 13*(1), 96-108.

Duflo, E. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review, 91*(4), 795-813. doi:10.1257/aer.91.4.795

Emmott, R., & Guarascio, F. (2016, 8 April 2021). Italy resists EU push for Russia sanctions over Aleppo bombings. *Reuters*. Retrieved from https://www.reuters.com/article/us-eu-summit-russia-idUSKCN12K2GK

European Commission. (2021). Carbon Border Adjustment Mechanism. Retrieved from https://ec.europa.eu/taxation_customs/green-taxation-0/carbon-border-adjustment-mechanism_en

European Parliament. (2022). EU carbon border adjustment mechanism: Implications for climate and competitiveness. Retrieved from https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2022)698889#:~:text=In%20parallel%2C%20a%20novel%20carbon,the%20weekly%20EU%20carbon%20price.

European Parliament and Council. (2008). Directive 2008/101/EC. *Official Journal of the European Union*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32008L0101

European Parliament and Council. (2013). Decision No 377/2013/EU. *Official Journal of the European Union*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32013D0377

European Parliament and Council. (2014). Regulation (EU) No 421/2014. *Official Journal of the European Union*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0421&from=EN

Farrell, H., & Newman, A. L. (2019). Weaponized Interdependence: How Global Economic Networks Shape State Coercion. *International Security, 44*(1), 42-79. doi:10.1162/isec_a_00351

Fricke, H. (2017). Identification Based on Difference-in-Differences Approaches with Multiple Treatments. *Oxford bulletin of economics and statistics, 79*(3), 426-433. doi:10.1111/obes.12178

Gleditsch, K. S. (2002). Expanded Trade and GDP Data. *The Journal of conflict resolution, 46*(5), 712-724. doi:10.1177/002200202236171

Gleditsch, K. S. (2013). *Updated GDP data to 2011*. Retrieved from: http://ksgleditsch.com/exptradegdp.html

Goldstein, J. L., Rivers, D., & Tomz, M. (2007). Institutions in International Relations: Understanding the Effects of the GATT and the WTO on World Trade. *International Organization, 61*(1), 37-67. doi:10.1017/S0020818307070014

Gomila, R. (2021). Logistic or linear? Estimating causal effects of experimental treatments on binary outcomes using regression analysis. *J Exp Psychol Gen, 150*(4), 700-709. doi:10.1037/xge0000920

Green, B. A. (2009). LESSONS FROM THE MONTREAL PROTOCOL: GUIDANCE FOR THE NEXT INTERNATIONAL CLIMATE CHANGE AGREEMENT. *Environmental Law, 39*(1), 253-283. Retrieved from http://www.jstor.org/stable/43266831

GreenAir. (2012). Countries opposing EU ETS to convene second meeting within weeks to discuss retaliatory action against Europe. Retrieved from https://archives.greenairnews.com/www.greenaironline.com/news46b6.html?viewStory=1417

Guzman, Andrew T., & Simmons, Beth A. (2005). Power Plays and Capacity Constraints: The Selection of Defendants in World Trade Organization Disputes. *The Journal of Legal Studies, 34*(2), 557-598. doi:10.1086/430767

Hedberg, M. (2018). The target strikes back: explaining countersanctions and Russia's strategy of differentiated retaliation. *Post-Soviet affairs, 34*(1), 35-54. doi:10.1080/1060586X.2018.1419623

Hellevik, O. (2007). Linear versus logistic regression when the dependent variable is a dichotomy. *Quality & quantity, 43*(1), 59-74. doi:10.1007/s11135-007-9077-3

Helm, D., Hepburn, C., & Ruta, G. (2012). Trade, climate change, and the political game theory of border carbon adjustments. *Oxford review of economic policy, 28*(2), 368-394. doi:10.1093/oxrep/grs013

Hepburn, C., Stern, N., & Stiglitz, J. E. (2020). "Carbon pricing" special issue in the European economic review. *Eur Econ Rev, 127*, 103440-103440. doi:10.1016/j.euroecorev.2020.103440

Hepher, T. (2012). China halts 10 more Airbus orders. *Reuters*. Retrieved from https://www.reuters.com/article/us-china-europe-ets-idUSBRE82E11620120315

Keohane, R. O. (2020). *International Institutions And State Power : Essays In International Relations Theory* (1st ed.): Routledge.

Keohane, R. O., & Nye, J. S. (1973). Power and interdependence. *Survival (London), 15*(4), 158-165. doi:10.1080/00396337308441409

Kherallah, M., & Beghin, J. (1998). U.S. trade threats: rhetoric or war? *U.S. trade threats: rhetoric or war?*(1), 15-29.

Klein, J. P., Goertz, G., & Diehl, P. F. (2006). The New Rivalry Dataset: Procedures and Patterns. *Journal of peace research, 43*(3), 331-348. doi:10.1177/0022343306063935

Lechner, M. (2011). The Estimation of Causal Effects by Difference-in-Difference Methods Estimation of Spatial Panels. *Foundations and Trends® in Econometrics, 4*, 165-224. doi:10.1561/0800000014

Mavroidis, P. C., Nordström, H., & Horn, H. (1999). *Is the use of the WTO dispute settlement system biased?* London: Centre for Economic Policy Research.

Milewicz, K. (2020). *Constitutionalizing world politics : the logic of democratic power and the unintended consequences of international treaty making*. Cambridge: Cambridge University Press.

Morgan, T. C., Bapat, N. A., & Kobayashi, Y. (2013). Threat and Imposition of Sanctions (TIES) Data 4.0 Users Manual. Retrieved from http://sanctions.web.unc.edu/

Morgan, T. C., Bapat, N. A., & Kobayashi, Y. (2014). Threat and imposition of economic sanctions 1945–2005: Updating the TIES dataset. *Conflict Management and Peace Science, 5*(31), 541–558.

Mueller, J. (1970). Presidential Popularity from Truman to Johnson. *The American Political Science Review, 1*(64), 18-34.

Multilateral Fund. (2022). Multilateral Fund for the Implementation of the Montreal Protocol. Retrieved from http://www.multilateralfund.org/default.aspx

Nordhaus, W. (2015). Climate Clubs: Overcoming Free-riding in International Climate Policy. *The American Economic Review, 105*(4), 1339-1370. doi:10.1257/aer.15000001

Nye, J. S. (2004). *Soft power : the means to success in world politics*. New York: PublicAffairs.

Pauer, S. (2019). *Border carbon adjustments in support of domestic climate policies : explaining the gap between theory and practice.* (Law). The University of British Columbia, Retrieved from https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0383398#downloadfiles

Peksen, D., & Jeong, J. M. (2021). Coercive Diplomacy and Economic Sanctions Reciprocity: Explaining Targets' Counter-Sanctions. *Defence and Peace Economics*, 1-17. doi:10.1080/10242694.2021.1919831

Reuters. (2012). Russia turns down extra overflights-airlines. *Reuters*. Retrieved from https://www.reuters.com/article/russia-airlines-overflight-idUSL5E8H88W620120608

Reuters. (2016, 8 April 2021). EU retreats from threat of Russia sanctions over Aleppo. Retrieved from https://www.reuters.com/article/mideast-crisis-syria-eu-idINKBN14111U

Russian Aviation. (2012). Joint declaration of the Moscow meeting on inclusion of international civil aviation in the Eu-ETS. Retrieved from https://www.ruaviation.com/docs/1/2012/2/22/50/

Sattler, T., Spilker, G., & Bernauer, T. (2014). Does WTO Dispute Settlement Enforce or Inform? *Brit. J. Polit. Sci, 44*(4), 877-902. doi:10.1017/S0007123413000136

Stacie, G. (2018). *When Right Makes Might : Rising Powers and World Order*. Ithaca: Ithaca: Cornell University Press.

UN Environment Programme. (2022). About Montreal Protocol. Retrieved from https://www.unep.org/ozonaction/who-we-are/about-montreal-protocol

United Nations. (2022). Montreal Protocol on Substances that Deplete the Ozone Layer. Retrieved from https://treaties.un.org/Pages/ViewDetails.aspx?src=IND&mtdsg_no=XXVII-2-a&chapter=27&clang=_en

van Aaken, A., & Simsek, B. (2021). Rewarding in International Law. *American Journal of International Law, 115*(2), 195-241. doi:10.1017/ajil.2021.2

World Bank. (2022). GDP (current US$). Retrieved from https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=EU-US-CN-IN-RU-CA

Zeng, K. (2004). *Trade threats, trade wars [electronic resource] : bargaining, retaliation, and American coercive diplomacy*. Ann Arbor: University of Michigan Press.

Appendix

A Instrumental variable model, including stage 1

	(1) Stage 1 DV: IO SUPPORT	(2) Stage 1 DV: IO X WEAK SENDER	(3) Stage 2 DV: TARGET AGAINST SENDER
IO SUPPORT			-0.031*
			(0.013)
IO X WEAK SENDER			-0.135**
			(0.050)
WEAK SENDER	0.232***	0.191***	0.122***
	(0.011)	(0.006)	(0.022)
INDUCEMENT TO COOPERATE	-0.051***	-0.002	-0.029***
	(0.007)	(0.004)	(0.009)
INDUCEMENT X WEAK SENDER	-0.085*	-0.136***	-0.171**
	(0.041)	(0.022)	(0.054)
IO MEMBERSHIP COUNT	0.002***	-0.000	
(instrument)	(0.000)	(0.000)	
IO MEMBERSHIP COUNT X WEAK SENDER	-0.001***	0.002***	
(instrument accounting for interaction)	(0.000)	(0.000)	
INITIAL ONGOING	-0.041***	0.001	0.042***
	(0.006)	(0.003)	(0.007)
EPISODE FEs	Yes	Yes	Yes
DECADE FEs	Yes	Yes	Yes
Observations	20,350	20,350	20,350
Number of EPISODEs	1,850	1,850	1,850

Standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05

B Full model (model 7) including IO MEMBERSHIP COUNT variable

VARIABLES	(1) Model 7 incl. IO MEMBERSHIP COUNT
IO SUPPORT	-0.025^
	(0.013)
WEAK SENDER	0.110***
	(0.026)
IO X WEAK SENDER	-0.108***
	(0.031)
INDUCEMENT TO COOPERATE	-0.029*
	(0.011)
INDUCEMENT X WEAK SENDER	-0.165***
	(0.046)
IO MEMBERSHIP COUNT	**-0.000**
	(0.000)
INITIAL EPISODE ONGOING	0.043***
	(0.012)
Constant	0.042***
	(0.006)
EPISODE FEs	Yes
DECADE FEs	Yes
Observations	20,350
Number of EPISODEs	1,850

Robust standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

C Instrumental variable model, including disputed issue variable and stage 1

VARIABLES	(1) Stage 1 DV: IO SUPPORT	(2) Stage 1 DV: IO X WEAK SENDER	(3) Stage 2 DV: TARGET AGAINST SENDER
IO SUPPORT			-0.026^
			(0.015)
IO X WEAK SENDER			-0.153**
			(0.050)
WEAK SENDER COALITION	0.149***	0.155***	0.131***
	(0.010)	(0.006)	(0.020)
INDUCEMENT TO COOPERATE	-0.142***	-0.045***	-0.006
	(0.008)	(0.004)	(0.011)
INDUCEMENT X WEAK SENDER	-0.034	-0.112***	-0.177**
	(0.038)	(0.021)	(0.054)
INITIAL ONGOING	-0.067***	0.024***	-0.004
	(0.012)	(0.007)	(0.016)
IO MEMBERSHIP COUNT	0.002***	-0.000	
(instrument)	(0.000)	(0.000)	
IO MEMBERSHIP COUNT X WEAK SENDER	-0.000	0.002***	
(instrument accounting for interaction)	(0.000)	(0.000)	
Disputed issue			
Contain political influence	0.073***	-0.063***	0.015
	(0.021)	(0.012)	(0.030)
Contain Military Behavior	0.351***	0.025**	0.020
	(0.015)	(0.008)	(0.022)
Destabilize Regime	0.051**	-0.013	0.028
	(0.016)	(0.009)	(0.023)
Release Citizens, Property, or Material	0.073***	-0.017	0.123***
	(0.018)	(0.010)	(0.025)
Solve Territorial Dispute	0.197***	0.032**	-0.002
	(0.019)	(0.011)	(0.028)
Deny Strategic Materials	0.071***	-0.019	0.030
	(0.018)	(0.010)	(0.026)
Retaliate for Alliance or Alignment Choice	0.265***	0.156***	0.021
	(0.014)	(0.008)	(0.022)
Improve Human Rights	0.257***	-0.002	0.032
	(0.014)	(0.007)	(0.020)
End Weapons/Materials Proliferation	-0.148***	-0.006	0.050
	(0.019)	(0.011)	(0.027)
Terminate Support of Non-State Actors	0.194***	-0.024	0.060
	(0.022)	(0.012)	(0.032)
Deter or Punish Drug Trafficking Practices	-0.021	0.031	0.025
	(0.031)	(0.017)	(0.044)
Improve Environmental Policies	0.038*	0.011	-0.018
	(0.017)	(0.010)	(0.025)
Trade Practices	-0.050***	-0.059***	0.056***
	(0.012)	(0.007)	(0.017)
Implement Economic Reform	0.509***	0.142***	0.031
	(0.016)	(0.009)	(0.024)
Other	-	-	-
(omitted as reference category)			
EPISODE FEs	Yes	Yes	Yes
DECADE FEs	Yes	Yes	Yes
Observations	20,350	20,350	20,350
Number of EPISODES	1,850	1,850	1,850

Standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, ^ p<0.1

D Prediction of inducements to cooperate (logit model)

	(1) BASIC	(2) FULL	(3) INCL. ISSUE	(4) **DV**: INTERNAL INDUCE.	(5) **DV**: INDUCE. WITHDRAW.	(6) **DV**: ADDITIONAL INDUCE.
WEAK SENDER COALITION	0.068***	0.062***	0.053***	0.307**	*Omitted due to lack of variation*	0.227***
	(0.000)	(0.000)	(0.000)	(0.025)		(0.005)
IO SUPPORT		0.726**	0.537***	0.438**	0.407***	7.403***
		(0.017)	(0.001)	(0.023)	(0.000)	(0.000)
TRADE LINKAGE		1.015***	1.013*	1.012	1.017***	0.990
		(0.001)	(0.078)	(0.116)	(0.001)	(0.408)
RIVALRY		0.476**	0.491**		0.555*	1.007
		(0.017)	(0.037)		(0.092)	(0.989)
Disputed issue						
Contain political influence		-	-			
(omitted as reference category)						
Contain Military Behavior			0.583			
			(0.269)			
Destabilize Regime			0.957			
			(0.926)			
Release Citizens, Property, or Material			0.690			
			(0.465)			
Solve Territorial Dispute			-			
Deny Strategic Materials			1.017			
			(0.976)			
Retaliate for Alliance or Alignment Choice			0.691			
			(0.417)			
Improve Human Rights			1.213			
			(0.663)			
End Weapons/Materials Proliferation			0.083***			
			(0.003)			
Terminate Support of Non-State Actors			0.129**			
			(0.018)			
Deter or Punish Drug Trafficking Practices			2.035			
			(0.309)			
Improve Environmental Policies			-			
Trade Practices			0.120***			
			(0.000)			
Implement Economic Reform			7.817***			
			(0.000)			
Other			3.434***			
			(0.007)			
Constant	0.389***	0.420***	0.749			
	(0.000)	(0.000)	(0.473)			
Observations	1,840	1,636	1,557			

Each cell entry is the exponential of the coefficient ($e^{coef.}$) which is the odds ratio. P-values are shown in parentheses.; *** p<0.001, ** p<0.01, * p<0.05, ^p<0.1

E Robustness tests: primary sender only, imposed only, threat only, controls

VARIABLES	(1) SUBSET: PRIMARY SENDER ONLY	(2) DV: IMPOSED ONLY	(3) DV: THREAT ONLY	(4) CONTROLS
IO SUPPORT	-0.039	-0.012	-0.017*	-0.024^
	(0.024)	(0.009)	(0.007)	(0.012)
WEAK SENDER	0.125***	0.066**	0.044*	0.103***
	(0.030)	(0.022)	(0.018)	(0.027)
IO X WEAK SENDER	-0.111**	-0.067**	-0.039^	-0.104**
	(0.043)	(0.025)	(0.021)	(0.035)
INDUCEMENT TO COOPERATE	-0.033^	-0.011	-0.018*	-0.032*
	(0.017)	(0.009)	(0.007)	(0.012)
INDUCEMENT X WEAK SENDER	-0.174***	-0.099**	-0.066**	-0.209***
	(0.052)	(0.035)	(0.024)	(0.049)
RIVALRY				0.040
				(0.030)
TRADE LINKAGE				0.000
				(0.001)
INITIAL EPISODE ONGOING	0.045***	0.017*	0.024***	0.035**
	(0.013)	(0.008)	(0.007)	(0.011)
Constant	0.037***	0.018***	0.024***	0.042***
	(0.008)	(0.005)	(0.004)	(0.008)
EPISODE FEs	Yes	Yes	Yes	Yes
DECADE FEs	Yes	Yes	Yes	Yes
Observations	13,629	20,350	20,350	17,285
Number of EPISODEs	1,239	1,850	1,850	1,633

Robust standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, ^p<0.1

F Robustness tests: varying definitions of senders' weakness

VARIABLES	(1) Weak sender: Target 150% GDP	(2) Weak sender: Target 200% GDP	(3) Weak sender: log GDP ratio	(4) Weak sender: material weakness score
IO SUPPORT	-0.027*	-0.032**	-0.087***	-0.028*
	(0.011)	(0.011)	(0.017)	(0.012)
WEAK SENDER COALITION (TARGET +150% GDP)	0.093***			
	(0.027)			
IO X WEAK SENDER COALITION (TARGET +150% GDP)	-0.132***			
	(0.031)			
WEAK SENDER COALITION (TARGET +200% GDP)		0.102***		
		(0.029)		
IO X WEAK SENDER COALITION (TARGET +200% GDP)		-0.121***		
		(0.035)		
WEAK SENDER COALITION (LOG GDP RATIO SENDER/TARGET)			0.014***	
			(0.003)	
IO X WEAK SENDER COALITION (LOG GDP RATIO SENDER/TARGET)			-0.015***	
			(0.003)	
WEAK SENDER COALITION (MATERIAL POWER SCORE TARGET>SENDER)				0.084***
				(0.023)
IO X WEAK SENDER COALITION (MATERIAL POWER SCORE TARGET>SENDER)				-0.088**
				(0.027)
INDUCEMENT TO COOPERATE	-0.039***	-0.038***	-0.017	-0.038**
	(0.011)	(0.011)	(0.013)	(0.012)
INITIAL EPISODE ONGOING	0.048***	0.049***	0.093***	0.041***
	(0.010)	(0.010)	(0.013)	(0.011)
Constant	0.042***	0.042***	0.043***	0.041***
	(0.006)	(0.006)	(0.006)	(0.007)
EPISODE FEs	Yes	Yes	Yes	Yes
DECADE FEs	Yes	Yes	Yes	Yes
Observations	19,639	19,639	19,639	18,661
Number of EPISODEs	1,821	1,821	1,821	1,717

Robust standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, ^p<0.1

G Robustness tests: different types of inducements to cooperate

VARIABLES	(1) Internal inducement / environmental issue	(2) Withdrawal of inducement	(3) Additional inducement	(4) All inducements
IO SUPPORT	-0.029*	-0.030*	-0.025*	-0.030*
	(0.011)	(0.012)	(0.012)	(0.012)
WEAK SENDER COALITION	0.111***	0.108***	0.113***	0.104***
	(0.025)	(0.026)	(0.025)	(0.026)
IO X WEAK SENDER COALITION	-0.102***	-0.100***	-0.105***	-0.102***
	(0.030)	(0.030)	(0.030)	(0.030)
ENVIRONMENTAL ISSUE (internal inducement)	-0.074**			-0.080**
	(0.024)			(0.025)
INDUCEMENT WITHDRAWAL (external inducement)		-0.020^		-0.024*
		(0.012)		(0.012)
ADDITIONAL INDUCEMENT (external inducement)			-0.022	-0.024
			(0.025)	(0.025)
INITIAL EPISODE ONGOING	0.036***	0.037***	0.033***	0.043***
	(0.009)	(0.010)	(0.009)	(0.010)
Constant	0.041***	0.042***	0.042***	0.041***
	(0.006)	(0.006)	(0.006)	(0.006)
EPISODE FEs	Yes	Yes	Yes	Yes
DECADE FEs	Yes	Yes	Yes	Yes
Observations	20,350	20,350	20,350	20,350
Number of EPISODEs	1,850	1,850	1,850	1,850

Robust standard errors in parentheses; *** p<0.001, ** p<0.01, * p<0.05, ^p<0.1

H Robustness tests: logit models

VARIABLES	(1) BASIC	(2) IO	(3) WEAK SENDER	(4) IO & WEAK SENDER	(5) INTERACTION	(6) INDUCE. TO COOPERA.	(7) INDUCE X WEAK SEN.	(8) FULL
IO SUPPORT		0.376***		0.372***	0.489***			0.471***
		(0.000)		(0.000)	(0.001)			(0.001)
WEAK SENDER			2.941***	3.036***	3.779***		2.595***	3.382***
			(0.000)	(0.000)	(0.000)		(0.000)	(0.000)
IO X WEAK SENDER					0.396**			0.385**
					(0.028)			(0.025)
INDUCEMENT TO COOPERATE						0.257***	0.366***	0.353***
						(0.000)	(0.000)	(0.000)
INDUCEMENT X WEAK SENDER							0.223	0.160
							(0.245)	(0.159)
INITIAL EPISODE ONGOING	2.511***	2.966***	1.742***	2.205***	2.083***	2.782***	2.069***	2.508***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.008***	0.010***	0.010***	0.011***	0.011***	0.010***	0.010***	0.011***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
EPISODE FEs	No	No	No	No	No	No	No	No
DECADE FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	20,922	20,922	20,350	20,350	20,350	20,922	20,350	20,350
Number of EPISODEs	1,902	1,902	1,850	1,850	1,850	1,902	1,850	1,850

Each cell entry is the exponential of the coefficient ($e^{coef.}$) which is the odds ratio. P-values are shown in parentheses.; *** p<0.001, ** p<0.01, * p<0.05, ^p<0.1