

Abstract *Scholars typically model the politics of global public goods or common pool resources as difficult collective action problems. Theories of international organization aim to explain how institutions can promote cooperation by solving the free-rider problem. Based on an analysis of a quintessential global collective action problem—international climate mitigation—this article challenges both this diagnosis of the problem and the concomitant institutional remedies. Important elements of climate mitigation exhibit three key features that depart from the canonical model: joint goods, preference heterogeneity, and increasing returns. The presence of these features creates the possibility for “catalytic cooperation.” Under such conditions, the chief barrier to cooperation is not the threat of free-riding, but the lack of incentive to act in the first place. States and other actors seek to solve this problem by creating “catalytic institutions” that work to shift actors’ preferences and strategies toward cooperative outcomes over time. While catalytic institutions can be seen in many areas of world politics, the 2015 Paris Agreement on climate change has put this logic of cooperation at its core, raising the possibility that similar catalytic institutions may facilitate cooperation in other areas of world politics characterized by analogous problem structures.*

To the extent collective action and free-riding dynamics do *not* drive climate policy—the premise of this special section—what are the implications for international cooperation? Traditionally, scholars have understood climate change as a tragedy of the commons, and international cooperation on mitigation as a kind of global Prisoners’ Dilemma.¹ They therefore advance solutions such as international institutions to monitor compliance, sanction defection, facilitate reciprocity, or assign property rights. But if we instead see climate policy as primarily driven by domestic distributional conflicts, are different strategies needed to promote cooperation, or is collective action simply not relevant?

This article aligns with Alkin and Mildemberger (this issue), arguing that the conventional wisdom mischaracterizes the nature of collective action for climate mitigation. However, it also demonstrates that cooperation remains important when certain conditions are met. Specifically, the article identifies features of cooperation problems that make “catalytic cooperation” possible, and shows that

* This article received valuable feedback and critique from Michaël Alkin, Liliana Andonova, Scott Barrett, Steven Bernstein, Joshua Busby, Jarred Finnegan, Fergus Green, Jessica Green, Matthew Hoffmann, Robert O. Keohane, Mareike Kleine, Matto Mildemberger, Michael Oppenheimer, Lauge Poulsen, Charles Roger, Anne-Marie Slaughter, Duncan Snidal, Harro van Asselt, Johannes Urpelainen, David Victor, and the participants of the London School of Economics International Political Economy Seminar, the Oxford International Political Economy seminar, the Nuffield College International Relations group, the London School of Economics Grantham Research Institute workshop, and the other authors of this special section, as well as several anonymous reviewers.

¹ For a discussion, see Alkin and Mildemberger, this issue.

these conditions apply to many areas of climate mitigation.² The article also explains how international institutions can promote cooperation in such circumstances not by solving the problems of free-riding and credible commitments, but by mobilizing incremental action that can shift the strategies and preferences of states and other actors over time.

The article illustrates its theoretical arguments through institutional shifts in the United Nations Framework Convention on Climate Change (UNFCCC). The 1997 Kyoto Protocol, following most other multilateral environmental agreements, took a “regulatory” approach, in which states negotiated a set of shared reductions targets. States are bound to report on their emissions and, should they miss the agreed target, face (minimal) sanctions. The 2015 Paris Agreement, instead, requires each country to put forward its own pledge, or “nationally determined contribution” (NDC). These pledges are then reviewed internationally and ratcheted up every five years. Paris also increases the role of cities, businesses, provinces and regions, and other sub- and non-state actors, creating institutions to recognize and orchestrate their climate action alongside the national pledges. Both these processes aim toward the long-term goal of limiting climate change to “well below” 2°C in this century, which requires effectively decarbonizing the world economy in the next decades.

Can this “pledge and review and ratchet” system work? Existing theories of international organization see the Paris Agreement as a positive but modest achievement for cooperation because it does little to solve free-rider problems and lacks robust enforcement provisions (Bang et al. 2016, Barrett & Dannenberg 2016, Keohane & Oppenheimer 2016, Sachs 2019). This article makes two theoretical arguments that invite a different interpretation.

First, it aims to correct a persistent misdiagnosis in international relations theory, which overemphasizes free-riding as the chief barrier to international cooperation around the global commons. Though the multiplayer Prisoners’ Dilemma represents *prima facie* a compelling interpretation of global climate mitigation, much of the observed political behavior and outcomes over the past two decades departs from its theoretical expectations (Alkin and Mildenberger, this issue). Three generalizable patterns emerge: climate policies provide private benefits as well as collective ones (“joint products”); the value actors place on mitigation, as well as the costs and benefits they face, vary enormously (“preference heterogeneity”); and action in the past lowers the costs and raises the benefits of action in the future (“increasing returns”).

To the extent these conditions apply, the catalytic cooperation model, which draws on other elements of cooperation theory, implies a very different set of barriers to, and strategies for, cooperation. If early movers reach a critical mass, and if past action has a large impact on subsequent action, cooperative action can become progressively self-reinforcing over time. This approach aligns with authors who see decarbonization as a problem of overcoming “lock in” through innovation and scaling (Bernstein & Hoffmann 2019, Farmer et al. 2019). This

² I am grateful to David Victor for suggesting this phrase.

article elaborates the implications for, and role of, international cooperation in this approach.

Second, the article introduces the concept of catalytic institutions, demonstrating how the Paris Agreement, and also international institutions in areas such as trade, human rights, and global health, attempt to support cooperation not by providing credibility, but through catalytic mechanisms that help build action over time by progressively shifting preferences. While it is by no means clear that the Paris Agreement will catalyze sufficient cooperation to avert dangerous changes in the Earth's climate, the article provides a theory to understand how it might.

The next section demonstrates how climate mitigation departs from conventional theoretical assumptions, identifying three conditions that define a catalytic model of cooperation. Section three examines the implications of this model for cooperation. The fourth section then elaborates the logic of catalytic institutions and shows how the Paris Agreement and other international institutions work to shift state preferences and strategies over time. In conclusion, the article identifies future research questions, implied by the model, which will need to be answered to understand prospects for catalytic cooperation in climate mitigation. It also considers how the problem structure of climate change may shift over time, and in what other issues areas catalytic institutions may support cooperation.

Three conditions that allow catalytic cooperation

Theories of collective action in political science, sociology, economics, and other fields focus on the incentive to contribute to collective action given the actions of others. Canonically, IR scholars focus on states, but because the theoretical set up below is more general, and because cities, businesses, and other such entities play an important role in national and global climate politics, I use the term "actors."

In the classic model the level of collective action provided is a production function P of the sum of all actors' contributions, A (i.e. the product of the number of contributors and the magnitude of their individual contributions) (Marwell & Oliver 1993). For a given actor i , the benefit of acting $g_i(a_i: A)$ is the value v_i it derives from the collective action $P(A)$, minus the cost c_i of its individual contribution a_i .

The classic model of collective action:

$$g_i(a_i: A) = v_i[P(A)] - c_i(a_i)$$

In this framework, the canonical interpretation of climate change is that no actor has an individual incentive to act absent some collective agreement; that is, on average $c_i(a_i) > v_i[P(A)]$. This situation creates a familiar and intuitive problem, a large-scale Prisoners' Dilemma (Hardin 1968, Barrett 2003, Sandler 2004). Bigger actors may have more to gain by acting, because they can do more to

affect the problem, but if no actor's contribution is by itself sufficient to gain an overall benefit that outweighs its cost of acting, collective action will not be forthcoming. Furthermore, even if some action should occur, an actor has little incentive to join in because it can more easily free-ride on others' efforts. Absent a credible commitment for all to act, no action is taken, and everyone ends up worse off.

But how accurately does the classic model of collective action capture the politics of climate mitigation or of other global collective action problems (Alkin and Mildenberger this issue)? A review of mitigation politics over the past decades provides evidence for three features of climate's problem structure that depart from strong assumptions of the Prisoners' Dilemma model:

1. *Joint products*: Contributions to collective action can yield "private" benefits for those that act in addition to public benefits shared by all.
2. *Preference heterogeneity*: Some actors value collective action highly, others little. Similarly, the costs and benefits of a given contribution to collective action vary across actors.
3. *Increasing returns*: Action in the past can reduce the cost and increase the benefit of action in the future, while also changing how actors value collective action.

Below I explain the logic of each revised assumption and demonstrate its empirical basis in the realm of climate mitigation. To the extent they apply, they shift the classic model of collective action described above to a "catalytic" model of collective action, which I define below.

Joint products

The classic model assumes that contributing to a collective good is costly to the contributor. In reality, of course, policy actions typically involve a mixture of costs and benefits to different actors. This means that collective goods are often "joint products" in that actors' contributions provide both a public good to the community and a private benefit to the contributor. Even seemingly pure public goods (e.g. funding basic research to create scientific knowledge) can come with private benefits of some kind (e.g. jobs for researchers). To the extent private benefits exist, we can expect at least some actors—what Olson (1971) termed a "privileged group"—to take at least some actions irrespective of others' behavior.

If the cost of a contribution tends to exceed the private benefit it provides the contributor, which we can separate out as a new term b_i , then the classic assumption is sensible. But the empirical record shows many cases in climate policy for which this assumption is too strong (Sandler 2004, p.43, Lipsy Forthcoming). After all, a country's net contribution to mitigation is almost never a single policy like a carbon tax, but rather a wide range of regulations and actions across nearly every sector including land-use policy, local zoning laws, transportation initiatives, development assistance, etc. Indeed, many actions relevant for mitigation may not be perceived primarily as climate policy per se.

Given this diversity, it is not surprising that many mitigation actions entail so-called “co-benefits”, including reducing local air pollution and improving human health, increasing energy security and reliability, developing new industrial sectors, preserving forests, reducing traffic, etc. Climate action can also bring narrower private benefits, such as material support for influential stakeholders (e.g. subsidies for farmers to make biofuels) or political benefits for politicians who face pro-climate constituencies. In some cases, these private benefits temper the cost of mitigation actions; in other cases, the other benefits are actors’ primary focus, and mitigation is an ancillary result. Even if it were true that, on average, most mitigation actions come with a net cost, some will not, and these “no cost” actions can allow actors to take significant steps even if they do not hold strong pro-climate views (Green 2018).

Preference heterogeneity

The classic model assumes relatively symmetric preferences across actors and actions, typically based in an intuitive but crude political economy model in which actors prioritize overall economic performance. While this simplifying assumption may be accurate on average, again, the empirical literature on climate change provides evidence of significant exceptions.

Actors face very different benefits and costs from mitigation policies. For example, countries with water resources like Sweden or Costa Rica have already nearly decarbonized electricity production through hydro-power. Moreover, the ‘natural’ variation actors face is exacerbated by the different decision-making institutions through which they form policy preferences, and the relative power of “winners” and “losers” of specific climate actions in decision-making processes. Comparative studies of national climate policy formation show how regime types and political-economic systems shape countries’ climate policies by valuing public goods differently, giving fossil fuel interests groups more or less power, or by giving more or less value to outcomes in the future (Dubash et al. 2013, Lachapelle & Paterson 2013, Bayer & Urpelainen 2016, Tobin 2017, Mildenberger 2020).

Expanding our understanding of actors’ preferences to consider non-material costs and benefits creates further possibilities to depart from the symmetric assumptions of the standard model. Even when there are net material costs to mitigation actions, it is not obvious that they always dominate other incentives, including ideational preferences for or against green policy. On average, decision makers certainly prefer to avoid expending money with little immediate return, but many decision-makers face pro-climate stakeholders, including citizens and voters, customers, and investors. Under these conditions, decision-makers may derive substantial political or financial benefits from acting.

Just as actors face sharply different costs and benefits of acting, so too do they value the benefits of collective action v_i differently. Climate impacts vary significantly across space and time, but it is common in the literature to assign a common “social cost of carbon” (an estimate of the impacts of future warming on welfare) and discount rate to the future benefits of mitigation. These simplifying

assumptions can obscure more than they clarify, since both impacts (Ricke et al. 2018) and time horizons vary significantly across actors. For small island states, arid countries near the equator, and coastal cities, the impacts are already severe and will become existential. For more temperate and inland areas, the changes will likely be slower and more moderate. Moreover, differential levels of economic development and state capacity will also mean that some actors will be able to adapt to climate disruptions much more effectively than others. Political institutions will also create variation in actors' discount factors, since different political systems value long- and short-term impacts differently (Jacobs 2016). For these reasons, we can expect v_i to vary dramatically across actors, with some motivated to act even if their action only generates a small level of collective good provision and others unlikely to act even if they could provide significant public goods unilaterally.³

Increasing returns

In the traditional set-up, action by others almost always dissuades an actor from acting. If others contribute, the actor can free-ride on their efforts. If others slack, the actor can scarcely hope to solve the problem without them.

Moreover, the model assumes, implicitly, that it is useful to conceptualize actors' preferences as if they do not change over time.⁴ In reality, some policy choices reinforce themselves through a variety of feedback loops that generate "increasing returns" to action over time (Pierson 2000). In this way prior action can shift actors' preferences toward further action over time. Of course, not all realms of climate policy exhibit increasing returns, and in some areas increasing returns have been too weak to overcome entrenched opposition (Stokes 2020). It is also possible to imagine that the opposite effect may occur, as first-movers harvest low-hanging fruit leaving only more costly actions for later-movers.

Ultimately, which dynamic dominates is an empirical question, but the literature highlights many examples of increasing returns at play (Urpelainen 2011, Levin et al. 2012, Urpelainen 2013, van der Ven et al. 2017). As Bernstein and Hoffmann note, decarbonization in one area can scale and entrench in other areas through normalization, capacity building, and coalition building (Bernstein & Hoffmann 2018). Material, informational, and normative mechanisms can be identified, affecting different parts of the model v_i , b_i , and c_i .

First, some climate action can affect the material costs of future action, c_i by changing technology and the economic systems around it. The literature on

³ Note that the magnitude of these effects is conditioned by what Sandler calls the "aggregation technology," which can be interpreted here as the nature of the production function P (Sandler 2004, p.68). When, as in mitigation, collective action is "summative," i.e., when every contribution helps create the public good, an actor can always get a bit of public benefit from its own contribution, no matter how small. But when a certain threshold of contributions must be reached before any meaningful collective benefit is created, actors, especially small ones, will not obtain any broader benefit v_i from unilateral action.

⁴ Historical institutionalist scholarship has noted the value of relaxing this last assumption, including the idea that changes in preferences may be endogenous to the decision to cooperate in the first place (Downs et al. 1998, Underdal et al. 2008).

socio-technical transitions shows how niche innovations by “pioneers” can diffuse to eventually reconfigure larger systems (Geels et al. 2017, Zenghelis et al. 2018, Victor et al. 2019). As technologies are developed and deployed, their costs descend down a “learning curve,” becoming cheaper as more R&D is conducted and as production and distribution systems “learn by doing” and reach economies of scale (Hale & Urpelainen 2015). Renewable energy technologies fit this pattern well. For example, the cost of photovoltaic cells dropped 75 percent from 2010-2015, and wind turbines 30-45 percent in the same period; both technologies are now at or below the cost of fossil fuel alternatives in many parts of the world, sharply altering the incentives for taking climate mitigation action (IRENA 2017). These cost reductions were only possible because of costly actions by first movers such as California, Denmark, and Germany (Sawin 2001, Breetz et al. 2018).

In addition to learning curves, many new technologies and business models demonstrate network effects; like telephones or email, the more people have them, the more useful they are, raising the private benefits b_i of such actions. Consider electric cars or solar rooftops. For the first movers, such products are very inconvenient because there are too few charging stations, maintenance technicians, or electric cables and pricing systems that allow homeowners to send power back to the grid. As market penetration increases, however, the enabling environment shifts and later adopters are well supported. Indeed, once network effects are strong enough, they may help to lock in new technologies as default options.

Second, early action can affect the political processes of preference formation for states and other actors by generating new constituencies for action (Urpelainen 2013, Meckling et al. 2015, van der Ven et al. 2017, Breetz et al. 2018). As new technologies emerge and grow, their producers and consumers develop a distributional interest in their continuance and expansion. At first these new interest groups are unlikely to be able to overcome established incumbents in political contestation. But in economic sectors or geographic regions where incumbents are relatively weak, the new entrants may thrive and, as action spreads, eventually acquire the size and clout to become politically competitive with incumbents in more and more jurisdictions and industries.

Third, prior action can also generate learning effects. Just as mitigation actions generate new technologies and business models that alter material costs, so too do they produce new knowledge about policy design and implementation through experimentation and diffusion, further reducing costs and enhancing benefits (Hoffmann 2011, Sabel & Victor 2017). This effect is significant because many areas of mitigation involve complex policy instruments. Emissions trading systems are a prominent example, where even highly competent regulators like the European Commission have struggled to make their credit allocation and exchange systems operate smoothly. For this reason, the Chinese government has drawn on significant international expertise in the development of its national emissions trading system, and proceeded incrementally by first experimenting at the provincial level. In this way, previous actions generate

epistemic resources that improve the ability of followers to emulate leaders (Biedenkopf et al. 2017).

Finally, over time growing actions may change norms around mitigation, altering not just costs and benefits, but also how actors value mitigation itself, v_i . International relations scholars have shown how “norm entrepreneurs” engage in “strategic social construction” by attempting to shift norms toward their policy goals (Finnemore & Sikkink 1998, p.895). In the realm of climate, Hoffmann and Bernstein argue that actions in one sphere, such as a city setting a climate target, can “normalize” low-carbon preferences in ways that spill over to other actors (Bernstein & Hoffmann 2018). These models follow a logic of increasing returns. As Finnemore and Sikkink argue, norms progress through a life-cycle from emergence, to a “norm cascade” in which they become widely followed in practice, to internalization, in which they are embedded in the beliefs and preferences of most actors. As more action takes place, the more this self-reinforcing logic applies. While decarbonization norms are not currently widespread, some areas of climate mitigation show evidence of norm cascade dynamics, such as divestment from fossil fuel companies (Green 2016).

To the extent these mechanism apply, prior actions increase the ratio of benefits b_i to costs c_i for subsequent actions, as well as the value actors place on mitigation v_i . Under these conditions actors’ preferences shift dynamically in ways not accounted for in the classic framework.

A catalytic model of cooperation

The above three modifications change the classic model of collective action into a catalytic model of collective action, which can be expressed as:

The catalytic model of collective action:

$$g_i(a_i; A) = v_i[P(A)] - c_i(a_i) + b_i(a_i)$$

Where:

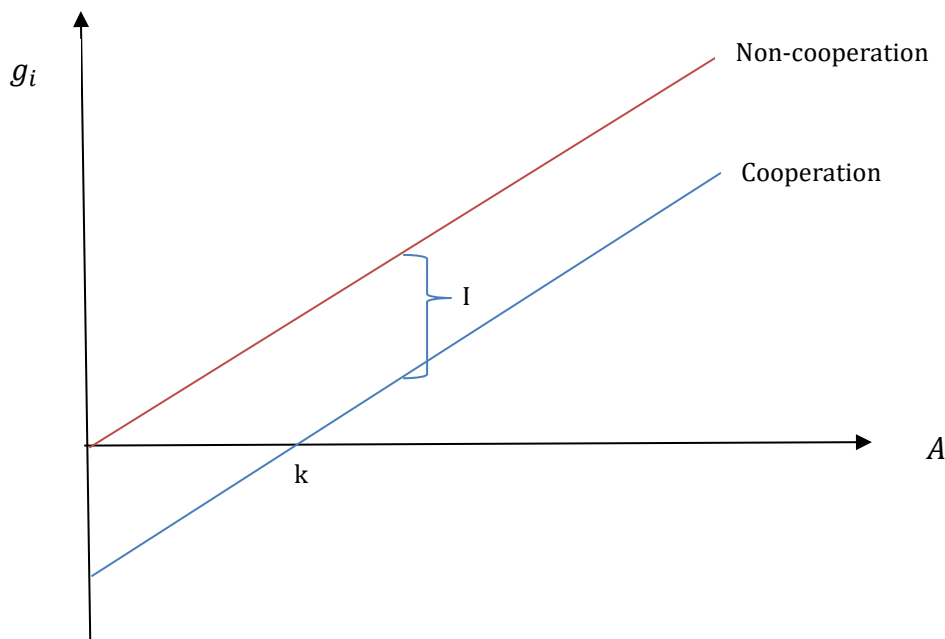
1. Private benefits exist, and in some cases $c_i(a_i) < b_i(a_i)$
2. Preferences vary such that, for some actors, $v_i[P(A)] > c_i(a_i)$ or, less restrictively, $v_i[P(A)] + b_i(a_i) > c_i(a_i)$
3. c_i , b_i , and v_i are partially functions of A such that
$$\begin{aligned} c_i &= C(A_{t-1}), C \text{ is a decreasing function} \\ b_i &= B(A_{t-1}), B \text{ is an increasing function} \\ v_i &= V(A_{t-1}), V \text{ is an increasing function} \end{aligned}$$

These three revisions to the core assumptions of the classic model, each grounded in an empirical analysis of climate politics, create the possibility for a very different form of collective action.

Implications for cooperation

The classic model suggests dim prospects for cooperation, as can be seen clearly in Schelling's classic representation (figure 1) (Schelling 1978, Snidal 1985). The vertical axis represents the pay-off to actor i from collective action as a function of the total amount of action, which is measured along the horizontal axis. Actor i may either cooperate or not cooperate, and the payoff for each choice is shown on a different curve. Because the non-cooperation curve is always higher than the cooperation curve, actor i will not act without an inducement, " I " at least as large as the distance between the cooperation curve and the non-cooperation curve. Inducements could come in the form of side-payments, sanctions, issue linkage, or, under repeated iterations, a credible pledge to unlock reciprocal cooperation from others.

Figure 1: Standard cooperation model



An exception occurs if a single actor or coalition of actors can contribute at least k , the point where cooperation and non-cooperation yield the same payoff. For example, a hegemonic state or coalition of great powers may have an interest in providing a public good, even if it means other states will free-ride on them, because the actions of the " k -group" are themselves sufficient to generate enough of the collective good to be worth the costs (Snidal 1985).

How does the catalytic model described above change the prospects for cooperation? Three implications follow from the revised assumptions.

Unilateral cooperation

Students of collective action have shown how joint products and preference heterogeneity can facilitate cooperation (Hardin 1982, Marwell & Oliver 1993). In the catalytic model, these characteristics make action strictly preferable to inaction for some actors in two ways:

1. The actor may value mitigation so much (v_i may be so high) that even a very low level of action A , potentially just the actor's own contribution a_i ,

- could produce a sufficient benefit to outweigh the cost $c_i(a_i)$ (heterogenous preferences).
2. The private benefit $b_i(a_i)$ of acting is higher than the cost $c_i(a_i)$ (joint products).

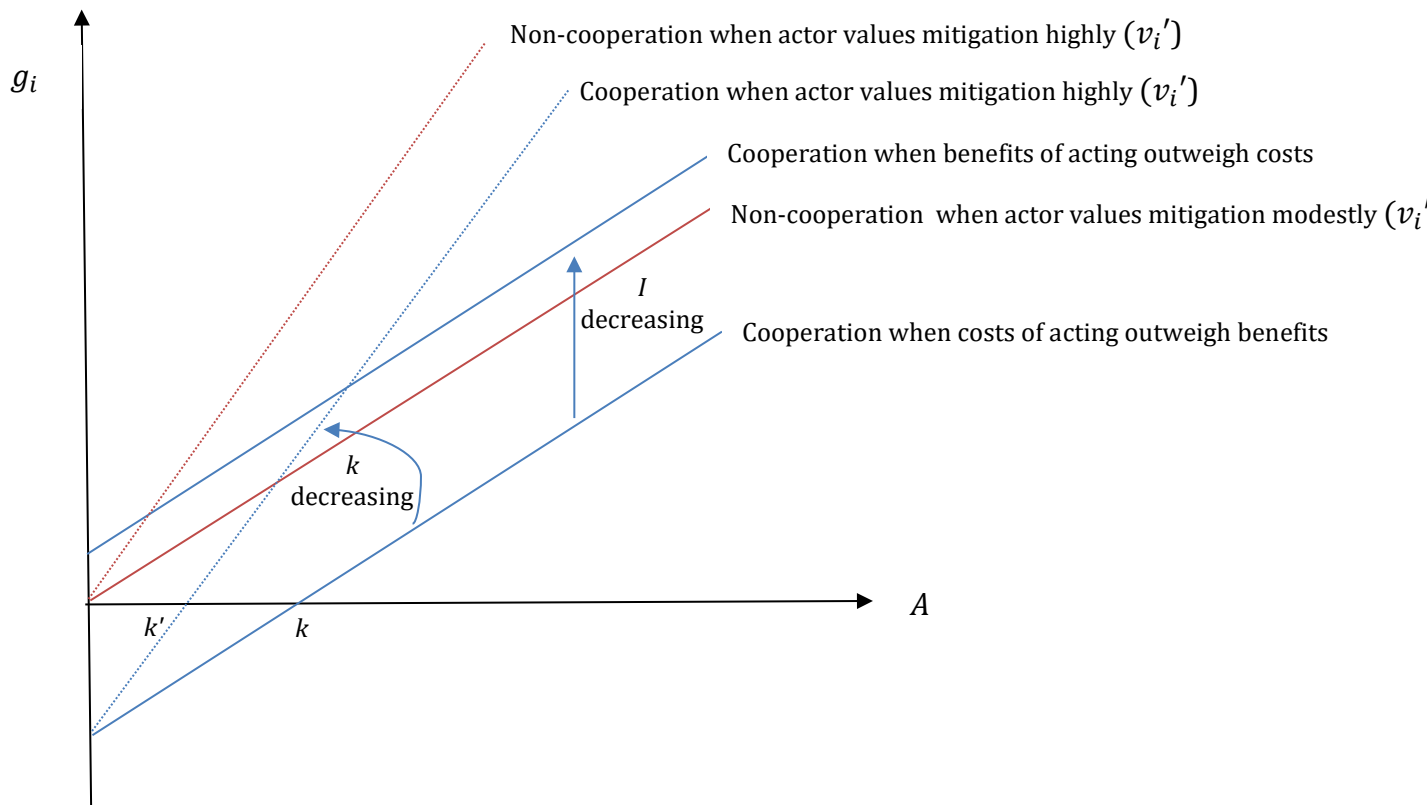
Combining these conditions, we can expect cooperation to be strictly preferable to non-cooperation whenever $v_i[P(A)] + b_i(a_i) > c_i(a_i)$. In Schelling's representation this "unilateral cooperation" can emerge as follows (figure 2).

First, as costs c_i fall and private benefits b_i rise, the cooperation curve shifts upward vertically, potentially falling above the non-cooperation curve. Even if costs remain higher than benefits, the catalytic model suggests that heterogenous preferences and joint products can reduce actors' inducement costs I .

Second, when actors value the benefits of mitigation more (say v_i is replaced by a higher value v_i'), the threshold for unilateral action k is lower, say k' . Valuing mitigation outcomes more increases the slope of both the cooperation and non-cooperation curves, rotating them counter-clockwise, which reduces the size of the k -group required to make cooperation attractive at a given ratio of costs to benefits.

The inverse situation, when v_i and b_i are low and c_i is high, is also of theoretical interest. Under these conditions, collective action is just not "worth it," meaning that even if a collective deal could be reached, it would not provide sufficient value to deter free-riding. In the realm of climate change, this situation is likely applicable at least in part to a number of significant actors, such as corrupt regimes with large oil reserves or fossil fuel-dependent companies focused on quarterly earnings reports.

Figure 2: How heterogenous preferences and joint products can induce unilateral cooperation

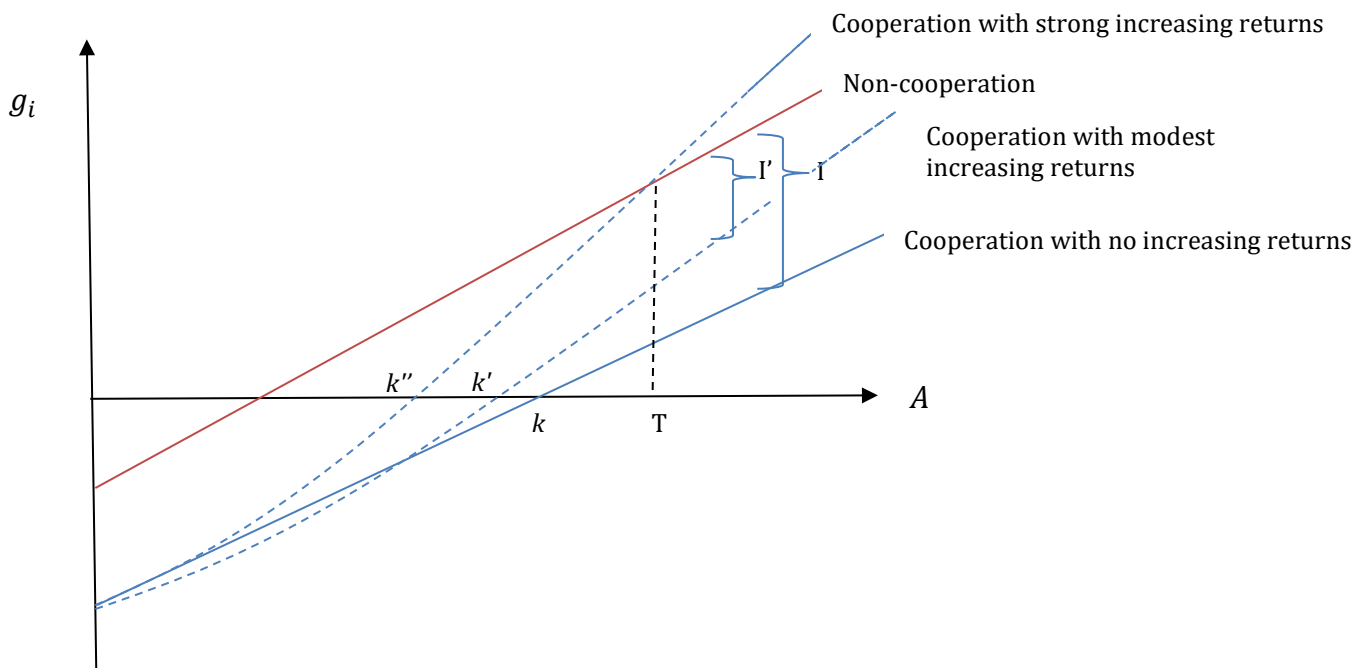


Increasing returns and tipping

In the catalytic model, increasing returns make v_i , b_i , and c_i partially endogenous to A , with the value of collective action and private benefits rising and costs falling as action expands. This change can be represented as an upward shift in the slope of the cooperation curve as action accumulates. In figure 3, as increasing returns strengthen, k shrinks to k' or even k'' , and I shrinks to I' .

To the extent increasing returns exist, a different pathway emerges for achieving cooperation: unilateral cooperators can become “first movers” who reduce the costs and increase the benefits of acting for followers and themselves. As Marwell and Oliver (1993, p. 88) note, “initial contributions lower the necessary interest for subsequent contributions.”

Figure 3: Catalytic cooperation with increasing returns



Similar ideas inform “cascade” or “threshold” models that have been used in political science to describe norm diffusion, protests and revolutions, racial segregation of neighborhoods, treaties that require a minimum critical mass to go into effect, and other phenomena, and have been used in experimental settings to analyze the Paris Agreement (Schelling 1971, Granovetter 1978, Kuran 1989, Finnemore & Sikkink 1998, Barrett 2003, Barrett & Dannenberg 2017). But these applications tend to describe coordination games of various kinds, in which actors are trying to converge on an equilibrium or shift from one equilibrium to another. For example, Schelling’s foundational study of neighborhood racial segregation showed how potentially even a small increase in the number of non-white residents could induce first the most anti-integration whites to move away, then the next-most racist ones, etc., potentially setting off a chain reaction that could change the racial composition of the entire neighborhood (Schelling 1972). Similarly, in Downs et al.’s (1998) model of evolutionary multilateralism, a small group of first movers is able to elicit greater cooperation over time by changing the strategic choices followers face. In all such models, actors’ preferences remain constant (though note that preference heterogeneity plays a key role); their choices change as others’ behavior changes.

In the catalytic model developed here, in contrast, actors’ very preferences alter as a consequence of their and other’s actions. If increasing returns are sufficiently strong, the cooperation curve may intersect and rise above the non-cooperation curve at a certain level of A . If this happens, a “tipping point” T occurs at which action becomes strictly preferable and self-reinforcing.

How the number of actors affects cooperation

In the classic collective action approach, a larger number of actors makes cooperation more difficult. If actors are relatively similar, a larger number of them means that each one's individual contribution will diminish relative to the whole, making it harder to assemble a k -group. Moreover, the transaction costs of negotiating and implementing a collective agreement to induce cooperation are higher (Olson 1971). For these reasons, a number of scholars have emphasized the value of "mini-lateral" institutions with smaller memberships (Eckersley 2012, Falkner 2016).

The catalytic cooperation model, in turn, helps us identify conditions under which increasing the number of actors can instead facilitate cooperation. Building on Olson and Hardin, Marwell and Oliver show that "when groups are heterogeneous and a good has high jointness of supply, a larger... group can have a smaller critical mass" (Marwell & Oliver 1993). This seemingly heterodox logic is deceptively simple. Assume, conservatively, that preferences are distributed normally across actors. Following the logic of joint products and heterogeneous preferences, raising the number of actors therefore increases the likelihood that at least a few actors will hold very pro-action preferences, because we become more likely to reach the "tails" of the distribution of preferences. Under a collective bargaining framework, this would do little to increase cooperation, because pro-action actors would, on average, be balanced by anti-action actors. But in a catalytic cooperation context where increasing returns apply, having a continuous chain of actors along the full preference spectrum—and especially at the pro-action extreme—is critical. This distribution of preferences increases the likelihood of at least one or more actors having a strong enough preference to act even when it is costly to do so.

Furthermore, having more actors in the system makes it easier for increasing returns to set off a chain reaction. As Granovetter's threshold model shows, a chain reaction can occur when actors' thresholds for action (k or T in the graphs above) are sufficiently close together, such that each new action is enough to reach the next-most pro-action actor's threshold (Granovetter 1978, Macy 1991). Assuming a sufficiently continuous distribution, increasing the number of actors means that it is less likely a large gap between thresholds will stall the effects of increasing returns.

Each of the three mechanisms described above shows how cooperation might emerge, given the assumptions of the catalytic cooperation model. While each individually highlights a potentially sufficient mechanism for collective action, they are also additive and complementary, together comprising a catalytic pathway to cooperation. This model can of course also potentially combine with "traditional" approaches to cooperation, such as collective agreements to ensure that the benefits of collective action are realized, sanctions against defectors, provision of exclusive benefits, etc.

Catalytic institutions

What role can international institutions play when commons problems are characterized by the catalytic cooperation model? The standard institutional solution of collective agreement, monitoring, and sanctions corresponds to the classic model of collective action. These features describe the basic contours of the Kyoto Protocol and many other agreements to regulate the global commons. However, as Schelling wrote following the Bush Administration's "un-signing" of the Kyoto Protocol, such an approach has not proven workable in climate mitigation (Schelling 2002).

Because the Paris Agreement allows actors to set their own targets and does not impose penalties for failing to meet them, the traditional theoretical lens gives reason to interpret Paris as a modest achievement. For example, Keohane and Oppenheimer (2016, p. 10) argue that Paris uses "discretion and vagueness" to make countries comfortable with making commitments (Keohane & Oppenheimer 2016, p.10). Future commitments will therefore be determined through domestic politics. Slightly more positively, Victor (2015) notes that "flexibility offers a way to get started and build confidence that, in time, will beget more confidence and a willingness to do more," but that "eventually a much more integrated global treaty will be needed to make major cuts in the greenhouse gases" (Victor 2015). Bang et al. (2016) are more critical, noting that "Paris does little to restructure states' incentives so as to avoid free riding," though they hope actors will become more pro-climate over time. Sachs (2019) instead finds little evidence to support the hope that peer pressure from other states or civil society will consistently shift countries toward greater emissions reductions. Barrett and Dannenberg (2016) present experimental evidence that pledge and review systems can help increase contributions, but not to the threshold required to reach the goal they have set. Following these arguments, we can see Paris is a positive but ultimately modest step.

The catalytic model offers a different interpretation of the Paris architecture. I do not make any claims here regarding why the treaty looks the way it does, nor do I consider it a 'perfect' example of a catalytic institution. Instead, this section uses key features of the Paris Agreement to illustrate the logic through which, when the conditions specified by the catalytic model are met, institutions can drive collective action by:

1. Stimulating first-movers and incremental action through flexibility
2. Iteration of commitments
3. Increasing the effect of prior action on subsequent action via:
 - a. Material transfers to alter future preferences and capacities
 - b. Transferring experiences to shape the informational environment
 - c. Normative goal-setting and benchmarking
 - d. Domestic constituency-building

Stimulating first-movers and incremental action through flexibility

In the catalytic cooperation model, the most important challenge is to reach a critical mass of action that can begin to generate increasing returns. Catalytic institutions therefore seek to stimulate first-movers to come forward and to encourage small steps even from recalcitrant actors.

By lowering the cost of action to what actors are willing to do, flexible commitments like NDCs can encourage early moves. If the bar for cooperation is too high, ie. a minimum a is required like a binding commitment required by Kyoto, these “small steps” might never occur (Victor 2011, Urpelainen 2013). Moreover, flexible commitments allow the most pro-mitigation actors to put forward ambitious commitments, instead of limiting themselves to a least common denominator negotiated outcome.

Similarly, the Paris Agreement seeks to spur first-movers amongst cities, businesses, and other actors (Hale 2016). Again, this is facilitated by the flexible nature of the Paris commitments, which allows for a variety of commitments and actions at varying degrees of ambition to be made. These actors add further to the mass of first-movers. Moreover as argued above, increasing the number of actors in this way makes it easier to create a chain reaction by shrinking the gaps between actors’ thresholds for action. A weakness of the voluntary approach, however, is that it does not necessarily elicit first-movers in sectors where, for example, the costs of acting remain high, meaning some segments of the economy (e.g. heavy industry) require further policy interventions to catalyze action (Victor et al. 2019).

Iteration of commitments

Even though flexible commitments capture only what actors are willing to do, with increasing returns what actors are willing to do changes over time. Catalytic institutions therefore create an ongoing process to record updates in actors’ preferences over time. By requiring new commitments every five years, the Paris Agreement allows increasing returns to be translated into new, stronger actions, which may then generate their own effects. A “one-off” commitment model would lack this ability to harvest the beneficial effects of past actions on subsequent actions over time.

As Macy notes in “cascade” models of collective action, it is much easier to achieve cooperation when choices are serial rather than parallel (Macy 1991). Indeed, iteration of commitment making is a common feature of international institutions. The global trade regime has developed through progressive trade rounds. More proximately, the “framework and protocol” approach to global commons issues like the ozone regime typically involves a series of increasingly stringent negotiated agreements, as was originally pursued in the climate regime (Sebenius 1991). While these regimes use collective agreements enforced with sanctions, they also have catalytic elements in that past cooperation alters future preferences. In trade, for example, the expansion of multinational companies enabled by early trade rounds re-shaped the domestic politics of economic openness in major economies by creating powerful new constituencies for integration (Bailey et al. 1997). Paris replicates this logic but with individual (for both countries and other actors), as opposed to negotiated, commitments, which allows it to capture updates in actors’ preferences more quickly and easily.

Increasing the effect of prior action on subsequent action

As discussed above, many of the mechanisms through which increasing returns accrue fall outside the realm of international institutions. For example, reductions in the price of clean technologies are transmitted through markets and social norms diffuse through a wide range of processes (Bernstein & Hoffmann 2019). However, international institutions can play a number of complementary roles to augment increasing returns .

a) Material transfers to alter future preferences and capacities

Many international agreements involve resources transfers, for example payments from one country to another to host a military base. Many such transfers are simply the inducement cost I needed to create cooperation. Transfers become catalytic when, above and beyond providing a direct side-payment, they increase the capability of the recipient to cooperate by affecting v_i , c_i and b_i . For example, a grant that helps pay for a renewable energy project reduces the one-off cost of cooperation. In contrast, a training and support package that improves a country's ability to run effective tendering processes for energy procurement could permanently reduce c_i , strengthening the country's preference for cooperation. Similarly, technical support that gives a country a better sense of what impacts it will suffer from climate change may increase the value v_i it places on mitigation efforts.

While the Paris Agreement did not contain any firm commitments to increase funding for developing countries (arguably the Kyoto Protocol was superior in this regard), and contributions to the Green Climate Fund have proven small, it did create a Paris Committee on Capacity Building that oversees and supports developing countries' ability to formulate and implement national climate policies. It remains to be seen, however, whether this process will actually generate resource flows that have a catalytic effect.

Instead, the Paris Agreement's greatest effect on material transfers may come indirectly. Many of the transnational networks of sub/non-state actors aim at peer-to-peer capacity building (Bulkeley et al. 2014). In addition, important bilateral donors and multilateral development banks are moving to align their portfolios with the goals of the Paris Agreement (MDB Paris Alignment Working Group 2018). These trends flow in part from the Paris Agreement's catalytic effect as a normative goal-setter (see below).

b) Sharing experiences to shape the informational environment

Many international institutions seek to generate and transfer information about state behavior in order to enhance compliance. As discussed in section three, such information is catalytic when it creates demonstration and learning effects that boost actors' ability and willingness to undertake mitigation actions, potentially reducing c_i and boosting b_i .

Several mechanisms in the Paris Agreement help to generate and diffuse these epistemic resources. First, flexible commitments, in addition to stimulating first-

movers and incremental action, can help drive the policy innovation through experimentation (Sabel & Victor 2017). As the difficult work of implementing these heterogeneous plans progresses, actors will gain useful experience of what works and what does not.

Second, catalytic institutions generate and diffuse information via review processes (Victor et al. 1998, Abbott 2017, Sabel & Victor 2017, Aldy 2018). In addition to assessing compliance, review processes allow experts to gather information about state behavior vis-à-vis an international obligation, generating and transmitting knowledge about how best to approach the problem (Chayes & Chayes 1995). Peer-to-peer transgovernmental networks have been shown to employ particularly influential versions of this type of review (Slaughter 2004), and have played an important role in the ozone regime and other global commons issues (Victor et al. 1998). The Paris Agreement includes review of individual countries' implementation of NDCs (Art. 13), review of aggregate progress toward the long-term goal in a Global Stock Take (Art. 14), and a "non-punitive" enforcement review designed to troubleshoot barriers to NDC implementation (Art. 15).

Third, a major goal of linking sub- and non-state actors to the intergovernmental regime is to promote learning. The number and diversity of sub- and non-state actors makes them excellent laboratories for climate policy, and many transnational initiatives explicitly promote information exchange (Roger et al. 2017). The Paris system tries to enhance the epistemic benefits of sub- and non-state climate action by creating structures to review and extract lessons from non-state actors including the NAZCA platform, the climate action events at COPs, and an annual Yearbook of Climate Action. Still, as Abbott notes, there is potential for the UNFCCC and other actors to play a more active role in enhancing the catalytic effect of these transnational elements of the regime (Abbott 2017).

c) Normative goal-setting and benchmarking

Section two noted how increasing action can normalize itself, potentially generating norm cascades, and therefore altering actors' preferences in favor of further action. Catalytic institutions seek to drive such processes through goal setting and benchmarking.

Countries often declare their collective intent to eradicate a disease, uphold human rights, or provide other global public goods, such as the Sustainable Development Goals. Article 2 of the Paris Agreement reaffirms countries' commitment to limit temperature change to "well below" 2C, aiming toward 1.5C, which, Article 4 notes, requires making sure no more GHGs are going into the atmosphere than are coming out by the middle of this century. As many observers have noted, absent concrete plans and accounting, goal-setting can be cheap talk or even intentionally dissembling (Downs et al. 1996). Nonetheless, scholars have identified various conditions under which, and mechanisms through which, goal-setting can affect political behavior (Biermann et al. 2017, Kanie et al. 2017).

First, a goal provides a focal point around which actors can converge. This mechanism is unlikely to sway actors who do not wish to cooperate, but it can enhance efficiency and enable benchmarking amongst actors guided by a logic of appropriateness, as first-movers are likely to be (Urpelainen 2013, Young 2017). In this regard, it is notable how the Paris Agreement's 1.5C and "net zero" goals have been adopted as a common target by a wide range of countries, businesses, sub-national governments, investors, and other actors (UNFCCC 2019).

Second, international goals can enhance the leverage of pro-cooperation constituencies in domestic politics, intra-firm deliberations, or other spheres of political contestation. To the extent states or other actors are sensitive about reputational critiques, explicit goal-setting raises the costs of non-compliance, and gives pro-compliance groups a "hook" for their arguments (Thomas 2001). For example, section two noted how decisionmakers sometimes act to satisfy pro-mitigation constituencies, who after 2015 would be unlikely to accept climate action that was not seen to be "Paris compatible."

Once a normative goal has been set, it becomes possible to review actors' progress toward that goal. Benchmarking systems—which grade states against some standard and compare their compliance to others—are now common tools in world politics (Kelley 2017). While some advocates proposed creating a grading system for NDCs in the Paris Agreement as part of the enhanced transparency framework, countries balked at exposing their "nationally determined" contributions to the collective judgement of their peers or others (suggesting the power of such ranking systems). Still, the Paris transparency framework requires countries to report on their progress toward their NDCs, and thus creates the conditions under which third parties can compare and rank national ambition (van Asselt 2016, Climate Action Tracker 2019).

d) Domestic constituency-building

Catalytic institutions also shift preferences by reaching into the black box of domestic politics and building new constituencies for greater climate action. While orchestration is widespread across many areas of world politics, the climate regime is perhaps unique in the extent to which policy entrepreneurs engaged in "webcraft" to orchestrate new actors to complement intergovernmental diplomacy (Hale & Roger 2014, Betsill et al. 2015, Hale 2016, Slaughter 2017). While transnational governance of climate change had been building over the course of the regime, in the lead up to Paris the UN Secretary-General, the UNFCCC Secretariat, and the Peruvian and French hosts of COP20 and COP21, respectively, took a much more purposeful approach, mobilizing dozens of initiatives that ultimately came to include over 10,000 actors. At Paris this orchestration function was then institutionalized in the UNFCCC process (Hale 2016). As described above, these efforts can have catalytic effects by influencing the process of national preference formation. The more sub- and non-state actors in a country take ambitious climate action, the more they increase the likelihood of both their peers and the national government adopting

strong climate policies as well (Urpelainen 2009, Bromley-Trujillo et al. 2016, Cao & Ward 2017).

Conclusion: where and when can catalytic institutions work?

The article has identified three features of climate mitigation—joint products, preference heterogeneity, and increasing returns—that depart from standard models of collective action. Together these characteristics, to the extent they apply, allow for catalytic cooperation. The argument is not that global climate mitigation has no resemblance to a large-scale Prisoners' Dilemma, but rather that choosing this, and only this, analytic model delivers a very partial understanding of global climate politics. In addition, the article has argued that catalytic institutions like those in the Paris Agreement can help drive international cooperation when conditions for catalytic cooperation obtain. Such institutions seek to initiate and stimulate early action, to iterate commitments over time, and to maximize the positive feedbacks from early action.

While the theory does not allow us to predict how effective the catalytic properties of the Paris Agreement will ultimately be, it points our attention to several factors on which its success will depend, identifying important areas for further research.

First, will enough first-movers appear? To generate catalytic cooperation, we need sufficient “unilateral cooperators” to emerge, either because they value cooperation highly (heterogenous preferences) or because the benefits of acting outweigh the costs (joint products). In first years since the Paris Agreement, a number of ambitious pledges have been made by small, progressive economies, vulnerable states, and European nations, as well as many significant sub- and non-state actors. But other large developed and emerging countries have either proceeded incrementally, stalled, or reversed (UNEP 2019). Indeed, the same elements of the model points our attention to the actors least likely to cooperate, those who face high decarbonization costs, who value mitigation little, and whose preferences are least likely to be shifted by increasing returns mechanisms (for example, fossil-dependent authoritarian regimes with relatively closed economies). To the extent these actors dominate, prospects for catalytic cooperation will remain bounded.

Second, to know whether this initial mass of first-movers is sufficient, we need to know the magnitude of increasing returns—material, informational, and ideational. For example, renewable energy has shown impressive growth in the power sector. Can we expect similar rapid scaling in transportation, land use, or heavy industry (Lafond et al. 2018)? Ideational shifts may be even more difficult to predict. For example, will rising civil society activism reshape climate politics globally, or will it be limited to a relatively small number of Western democracies? Will changes in dietary norms scale up to meaningfully reduce deforestation? Empirical answers to these questions are needed to assess the potential for catalytic cooperation to proceed in climate change.

The model also leads us to ask how its assumptions might change over time. Catalytic cooperation likely has more analytic power during earlier phases of cooperation rather than later ones, since it is premised on partially endogenous shifts in preferences that can change the nature of the cooperation problem. Various outcomes are possible in the medium- to long-term. Most optimistically, every actor of consequence could “tip” and pro-climate preferences diffuse globally. Under such conditions, the problem structure would change into a coordination game, and we would expect actors to create institutions to provide focal points and facilitate implementation (Barrett 2003).

Alternatively, or in a previous phase, we may see particularly recalcitrant actors fail to shift, even as the majority of others do. This could lead to a problem structure not dissimilar to that of nuclear proliferation in which a few “rouges” impose externalities on others. In this conflictive scenario, the bulk of states could have an incentive to develop more coercive club institutions to exclude and penalize the holdout emitters, for example through carbon tariffs. In this way, catalytic cooperation may lead to the conditions needed for the emergence of a climate club that could use negative incentives to enforce cooperation (Nordhaus 2015).

Of course, it may also be the case that the Paris Agreement will fail to catalyze sufficient action to force a critical mass of actors to raise their mitigation ambition (Sachs 2019). Under this scenario, the climate regime will look much as it did after Copenhagen, characterized by a flexible multilateral process and increasing pluralism. As the polycentric logic suggests, this will involve many different approaches across different actors reflecting their individual preferences and strategies, but may not add up to a global solution (Jordan et al. 2018).

Finally, could such a catalytic approach work in other issue areas? A number of international cooperation dilemmas exhibit joint products, preference heterogeneity, and increasing returns. As section three discussed, these features imply conditions under which catalytic cooperation becomes possible:

1. When there is high variation in actors’ preferences, with at least some willing to act unilaterally (in other words, on average, the costs and benefits of free-riding are low).
2. Prior action alters the costs, benefits, and preference formation processes around subsequent actions, potentially creating tipping points
3. Many different actors can affect the problem

Clearly, many areas of world politics do not fit this description. For example, the cost/benefit of free-riding in the realm of nuclear proliferation is high. If one country agrees to take in a number of refugees, it does not lower the cost for other countries to take in additional refugees in the future. And while many actors are affected by trade policies, only sovereign states can raise or lower tariffs.

Still, many other areas of world politics, and especially the realms of development, environment, and social welfare, satisfy these conditions. For

example, of the 17 Sustainable Development Goals, areas like health, gender equality, water and sanitation, food security, energy, urbanization, and responsible consumption and production all are relatively insulated from free-riding concerns and affected by a wide range of actors with highly variable preferences. These tend to be, like climate change, classic “intermestic” issues that have domestic political dimensions but also spill across borders. We should therefore expect catalytic cooperation and catalytic institutions to continue to play an important role in global governance.

Biography

Thomas Hale is Associate Professor of Global Public Policy at the Blavatnik School of Government, University of Oxford. His research explores how we can manage transnational problems effectively and fairly. He holds a PhD in Politics from Princeton University, a masters degree in Global Politics from the London School of Economics, and an AB in public policy from Princeton’s Woodrow Wilson School. His books include *Beyond Gridlock* (Polity 2017), *Between Interests and Law: The Politics of Transnational Commercial Disputes* (Cambridge 2015), *Transnational Climate Change Governance* (Cambridge 2014), and *Gridlock: Why Global Cooperation Is Failing when We Need It Most* (Polity 2013).

Bibliography

- Abbott, K. W. (2017). "Orchestrating experimentation in non-state environmental commitments." *Environmental Politics* **26**(4): 738-763.
- Aldy, J. E. (2018). Policy Surveillance: Its Role in Monitoring, Reporting, Evaluating and Learning. *Governing Climate Change: Polycentricity in Action*. A. Jordan, D. Huitema, H. van Asselt and J. Forster. Cambridge, Cambridge University Press.
- Bailey, M. A., J. L. Goldstein and B. R. Weingast (1997). "The Institutional Roots of American Trade Policy: Politics, Coalitions, and International Trade." *World Politics* **49**(3): 309-338.
- Bang, G., J. Hovi and T. Skodvin (2016). "The Paris Agreement: Short-term and Long-term Effectiveness." *Politics and Governance* **4**(3): 209-218.
- Barrett, S. (2003). *Environment and Statecraft*. Oxford, Oxford University Press.
- Barrett, S. and A. Dannenberg (2016). "An Experimental Investigation into 'Pledge and Review' in Climate Negotiations." *Climatic Change* **138**: 339-351.
- Barrett, S. and A. Dannenberg (2017). "Tipping versus Cooperating to Supply a Public Good." *Journal of the European Economic Association* **15**(4): 910-941.
- Bayer, P. and J. Urpelainen (2016). "It Is All about Political Incentives: Democracy and the Renewable Feed-in Tariff." *Journal of Politics* **78**(2): 603-619.
- Bernstein, S. and M. Hoffmann (2018). "The politics of decarbonization and the catalytic impact of subnational climate experiments." *Policy Sciences* **51**(2): 189-211.
- Bernstein, S. and M. Hoffmann (2019). "Climate politics, metaphors and the fractal carbon trap." *Nature Climate Change* **9**(12): 919-925.
- Betsill, M., N. K. Dubash, M. Paterson, H. van Asselt, A. Vihma and H. Winkler (2015). "Building Productive Links between the UNFCCC and the Broader Global Climate Governance Landscape." *Global Environmental Politics* **15**(2): 1-10.

- Biedenkopf, K., P. Müller, P. Slominski and J. Wettestad (2017). "A Global Turn to Greenhouse Gas Emissions Trading? Experiments, Actors, and Diffusion." Global Environmental Politics **17**(3): 1-11.
- Biermann, F., N. Kanie and R. E. Kim (2017). "Global governance by goal-setting: the novel approach of the UN Sustainable Development Goals." Current Opinion in Environmental Sustainability **26-27**: 26-31.
- Breetz, H., M. Mildemberger and L. Stokes (2018). "The political logics of clean energy transitions." Business and Politics **20**(4): 492-522.
- Bromley-Trujillo, R., J. S. Butler, J. Poe and W. Davis (2016). "The Spreading of Innovation: State Adoptions of Energy and Climate Policy." Review of Policy Research **33**(5): 544-565.
- Bulkeley, H., L. B. Andonova, M. Betsill, D. Compagnon, T. Hale, M. Hoffmann, P. Newell, M. Paterson, C. Roger and S. D. VanDeveer (2014). Transnational Climate Change Governance. Cambridge, Cambridge University Press.
- Cao, X. and H. Ward (2017). "Transnational Climate Governance Networks and Domestic Regulatory Action." International Interactions **43**(1): 76-102.
- Chayes, A. and A. H. Chayes (1995). The New Sovereignty: Compliance with International Regulatory Agreements. Cambridge, Harvard University Press.
- Climate Action Tracker (2019). Warming Projections Global Update, NewClimate Institute.
- Downs, G., D. M. Rocke and P. N. Barsoom (1996). "Is the Good News about Compliance Good News about Cooperation?" International Organization **50**(2): 379-406.
- Downs, G., D. M. Rocke and P. N. Barsoom (1998). "Managing the Evolution of Multilateralism." International Organization **52**(2): 397-419.
- Dubash, N. K., M. Hagemann, N. Hohne and P. Upadhyaya (2013). "Developments in national climate change mitigation legislation and strategy." Climate Policy **13**(6): 649-664.
- Eckersley, R. (2012). "Moving Forward in the Climate Negotiations: Multilateralism or Minilateralism?" Global Environmental Politics **12**(2): 24-42.
- Falkner, R. (2016). "A Minilateral Solution for Global Climate Change? On Bargaining Efficiency, Club Benefits, and International Legitimacy." Perspectives on Politics **14**(1): 87-101.
- Farmer, J. D., C. Hepburn, M. C. Ives, T. Hale, T. Wetzler, P. Mealy, R. Rafaty, S. Srivastav and R. Way (2019). "Sensitive intervention points in the post-carbon transition." Science **364**(6436): 132-134.
- Finnemore, M. and K. Sikkink (1998). "International Norm Dynamics and Political Change." International Organization **52**(4): 887-917.
- Geels, F. W., B. K. Sovacool, T. Schwanen and S. Sorrell (2017). "Sociotechnical transitions for deep decarbonization." Science **357**(6357): 1242.
- Granovetter, M. (1978). "Threshold Models of Collective Behavior." The American Journal of Sociology **83**(6): 1978.
- Green, F. (2016). "Anti-fossil fuel norms." Forthcoming.
- Green, F. (2018). "Anti-fossil fuel norms." Climatic Change **150**(1): 103-116.
- Hale, T. (2016). "'All Hands on Deck': The Paris Agreement and Nonstate Climate Action." Global Environmental Politics **16**(3): 12-22.
- Hale, T. and C. Roger (2014). "Orchestration and Transnational Climate Governance." Review of International Organizations **9**(1): 59-82.

- Hale, T. and J. Urpelainen (2015). "When and how can unilateral policies promote the international diffusion of environmental policies and clean technology?" Journal of Theoretical Politics **27**(2): 177 - 205.
- Hardin, G. (1968). "The Tragedy of the Commons." Science **162**(1243).
- Hardin, R. (1982). Collective Action. Baltimore, Johns Hopkins University Press.
- Hoffmann, M. (2011). Climate Governance at the Crossroads: Experimenting with a Global Response after Kyoto. Oxford, Oxford University Press.
- IRENA. (2017). "'Wind' and 'Solar'." Retrieved December 27, 2017, from www.irena.org.
- Jacobs, A. M. (2016). "Policy Making for the Long Term in Advanced Democracies." Annual Review of Political Science **19**(1): 433-454.
- Jordan, A., D. Huitema, H. van Asselt and J. Forster (2018). Governing Climate Change: Polycentricity in Action. Cambridge, Cambridge University Press.
- Kanie, N., S. Bernstein, F. Biermann and P. M. Haas (2017). Introduction Global Governance through Goal Setting. Governing through Goals. N. Kanie and F. Biermann, Mit Press: 1-28.
- Kelley, J. G. (2017). Scorecard Diplomacy: Grading States to Influence their Reputation and Behavior. Cambridge, Cambridge University Press.
- Keohane, R. O. and M. Oppenheimer (2016). "Paris: Beyond the Climate Dead End through Pledge and Review?" Politics and Governance **4**(3): 142-151.
- Kuran, T. (1989). "Sparks and Prairie Fires: A Theory of Unanticipated Political Revolution." Public Choice **61**(1): 41-74.
- Lachapelle, E. and M. Paterson (2013). "Drivers of national climate policy." Climate Policy **13**(5): 547-571.
- Lafond, F., A. G. Bailey, J. D. Bakker, D. Rebois, R. Zadourian, P. McSharry and J. D. Farmer (2018). "How well do experience curves predict technological progress? A method for making distributional forecasts." Technological Forecasting and Social Change **128**: 104-117.
- Levin, K., B. Cahsore, S. Bernstein and G. Auld (2012). "Overcoming the tragedy of super wicked problems: constraining our future selves to ameliorate global climate change." Policy Sci **45**(123-152).
- Lipsky, P. Y. (Forthcoming). The Institutional Politics of Energy and Climate Change.
- Macy, M. W. (1991). "Chains of Cooperation: Threshold Effects in Collective Action." American Sociological Review **56**(6): 730-747.
- Marwell, G. and P. Oliver (1993). The Critical Mass in Collective Action. Cambridge, Cambridge University Press.
- MDB Paris Alignment Working Group (2018). The MDBs' alignment approach to the objectives of the Paris Agreement: working together to catalyse low-emissions and climate-resilient development World Bank.
- Meckling, J., N. Kelsey, E. Biber and J. Zysman (2015). "Winning coalitions for climate policy." Science **349**(6253): 1170.
- Mildenberger, M. (2020). Carbon Captured: How Business and Labor Control Climate Politics. Cambridge, MIT Press.
- Nordhaus, W. (2015). "Climate Clubs: Overcoming Free-riding in International Climate Policy." American Economic Review **105**(4): 1339-1370.
- Olson, M. (1971). The Logic of Collective Action: Public Goods and the Theory of Groups. Cambridge, Harvard University Press.
- Pierson, P. (2000). "Increasing Returns, Path Dependence, and the Study of Politics." American Political Science Review **94**(2): 251-267.

- Ricke, K., L. Drouet, K. Caldeira and M. Tavoni (2018). "Country-level social cost of carbon." Nature Climate Change **8**(10): 895-900.
- Roger, C., T. Hale and L. B. Andonova (2017). "The Comparative Politics of Transnational Climate Governance." International Interactions **43**(1): 1-25.
- Sabel, C. F. and D. G. Victor (2017). "Governing global problems under uncertainty: making bottom-up climate policy work." Climatic Change **144**(1): 15-27.
- Sachs, L. E. (2019). "The Paris Agreement in the 2020s: Breakdown or Breakup?" Ecology Law Quarterly **46**(1).
- Sandler, T. (2004). Global Collective Action. Cambridge, Cambridge University Press.
- Sawin, J. L. (2001). The role of government in the development and diffusion of renewable energy technologies: wind power in the United States, California, Denmark and Germany, 1970–2000. PhD, Tufts University.
- Schelling, T. C. (1971). "Dynamic Models of Segregation." The Journal of Mathematical Sociology **1**(2): 143-186.
- Schelling, T. C. (1972). The Process of Residential Segregation: Neighborhood Tipping. Racial Discrimination in Economic Life. A. H. Pascal. Lexington, Lexington Books.
- Schelling, T. C. (1978). Micromotives and Macrobehavior. New York, Norton.
- Schelling, T. C. (2002). "What Makes Greenhouse Sense?" Foreign Affairs **81**(3): 2-9.
- Sebenius, J. K. (1991). "Designing Negotiations Toward a New Regime: The Case of Global Warming." International Security **15**(4): 110-148.
- Slaughter, A.-M. (2004). A New World Order. Princeton, Princeton University Press.
- Slaughter, A.-M. (2017). The Chessboard and the Web: Strategies of Connection in a Dangerous World. New Haven, Yale University Press.
- Snidal, D. (1985). "The Limits of Hegemonic Stability Theory." International Organization **39**(4): 579-614.
- Stokes, L. C. (2020). Short Circuiting Policy: Interest Groups and the Battle Over Clean Energy and Climate Policy in the American States. Oxford, Oxford University Press.
- Thomas, D. C. (2001). The Helsinki Effect: International Norms, Human Rights, and the Demise of Communism. Princeton, Princeton University Press.
- Tobin, P. (2017). "Leaders and Laggards: Climate Policy Ambition in Developed States." Global Environmental Politics **17**(4): 28-47.
- Underdal, A., S. Kallbenkken and J. Hovi (2008). "Exploring the Dynamics of International Cooperation." International Studies Association Annual Convention.
- UNEP (2019). The Emissions Gap Report 2019. Nairobi, United Nations Environment Programme.
- UNFCCC (2019). Climate Ambition Alliance: Nations Renew their Push to Upscale Action by 2020 and Achieve Net Zero CO2 Emissions by 2050, UNFCCC.
- Urpelainen, J. (2009). "Explaining the Schwarzenegger Phenomenon: Local Frontrunners in Climate Policy." Global Environmental Politics **9**(3): 82-105.
- Urpelainen, J. (2011). "A California Effect for International Environmental Externalities." International Interactions **37**(2): 170-189.
- Urpelainen, J. (2013). "A model of dynamic climate governance: dream big, win small." International Environmental Agreements **13**: 107-125.
- van Asselt, H. (2016). "The Role of Non-State Actors in Reviewing Ambition, Implementation, and Compliance under the Paris Agreement." Climate Law **6**(1): 91-108.

- van der Ven, H., S. Bernstein and M. Hoffmann (2017). "Valuing the Contributions of Nonstate and Subnational Actors to Climate Governance." Global Environmental Politics **17**(1): 1-20.
- Victor, D. G. (2011). Global Warming Gridlock: Creating More Effective Strategies for Protecting the Planet. Cambridge, Cambridge University Press.
- Victor, D. G. (2015). "Why Paris Worked: A Different Approach to Climate Diplomacy,," Yale360 2015.
- Victor, D. G., F. W. Geels and S. Sharpe (2019). Accelerating the Low Carbon Transition: The Case for Stronger, More Targeted and Coordinated International Action. London, Energy Transitions Commission
- Victor, D. G., K. Raustiala and E. B. Skolnikoff (1998). The Implementation and Effectiveness of International Environmental Commitments. Cambridge, MIT Press.
- Young, O. R. (2017). Conceptualization: Goal Setting as a Strategy for Earth System Governance. Governing through Goals: Sustainable Development Goals as Governance Innovation. N. Kanie and F. Biermann. Governing through Goals: Sustainable Development Goals as Governance Innovation, MIT Press.
- Zenghelis, D., R. Fouquet and R. Hippe (2018). Stranded assets: then and now. Stranded Assets and the Environment: Risk, Resilience and Opportunity. B. Caldecott. London, Taylor and Francis.