CUADERNOS ECONÓMICOS ICE INFORMACIÓN COMERCIAL ESPAÑOLA

Behavior, decisions and ecological transition: experimental approaches with policy implications

Comportamiento, decisiones y transición ecológica: aproximaciones experimentales con implicaciones para las políticas

Alberto Antonioni
Antonio Cabrales
Francesca Lipari
Anxo Sánchez
Universidad Carlos III de Madrid

Abstract

One key aspect of the ecological transition has to do with individuals' and collective behavior and its impact on climate change and decarbonization. We will describe how these questions can be studied by proper experimental designs by focusing on two examples: the implementation of Nordhaus's climate club idea for making countries contribute to climate change mitigation, and the interplay between the perception of the risk of a climatic catastrophe and the social norms arising from, and influencing, individual behavior. We will draw conclusions of the outcome of the experiments that can be relevant for policy making.

Keywords: *climate clubs, behavioral experiments, social norms.*
JEL codes: *D9, Q5.*

Resumen

Un aspecto clave de la transición ecológica tiene que ver con el comportamiento individual y colectivo y su impacto en el cambio climático y la descarbonización. Describimos cómo pueden estudiarse estas cuestiones mediante diseños experimentales adecuados, centrándonos en dos ejemplos: la aplicación de la idea de los clubes climáticos de Nordhaus para hacer que los países contribuyan a la mitigación del cambio climático, y la interacción entre la percepción del riesgo de una catástrofe climática y las normas sociales que se derivan de, y que influyen en, el comportamiento individual. En las conclusiones discutimos cómo este tipo de experimentos pueden ser relevantes para la elaboración de políticas.

Palabras clave: *clubs climáticos, experimentos conductuales, normas sociales.*

DOI: https://doi.org/10.32796/cice.2022.104.7505
Recibido: Abril 2022 · Aceptado: Septiembre 2022

1. Introduction

According to Masson-Delmotte et al. (2021), each of the last four decades has been successively warmer than any decade that preceded it since 1850, and global surface temperature in the period 2001-2020 was around 1 °C higher than 1850-1900. Only in the last few years a clear majority, ranging from 83% in USA to 95% in Germany in the nine countries surveyed by Eichhorn et al. (2020), of Europeans and US-Americans have finally become aware that the climate is changing, and also that human activity is an important cause of climate change (from 79% in USA to 90% in Italy). In Spain, 97% of people surveyed by Lázaro Touza et al. (2019) agreed that climate change exists, and 92% of them agreed that it is caused by human activity.

In the UN Climate Change Conference (COP21) in Paris, in 2015, the participant countries reached a historic agreement, an internationally binding treaty to address this crisis. The agreement established, among other things, a set of *Nationally Determined Contributions* that committed the countries to a fixed level of abatement of emissions. They also committed to establish long-term strategies towards achieving net zero emissions. Interestingly, from this point of view, the latest Intergovernmental Panel on Climate Change (IPCC, 2022) report claims that "Collective action and strengthened networked collaboration, more inclusive governance, spatial planning and risk-sensitive infrastructure delivery will contribute to reducing risks." That is, there exists a growing realization that social issues are a key consideration to deal with this problem.

Such is the approach we take in this paper. Climate change is a serious collective action problem. No one single individual or even country, however powerful, can deal with it alone. This means that strategic interaction is inevitable when addressing mitigation efforts. Another important characteristic of the problem is that we cannot really do many experiments in the field. As the saying goes: "there is no planet B." Both strategic interaction and the difficulty of doing field experiments suggests that laboratory experimentation with games is one of the few good empirical models to understand factors affecting behavioral mitigation efforts.

The usefulness of experiments arises from different reasons. First, the orthodox economics models that regulators and policymakers use to understand the effect of various mitigation policies usually ignore that humans are boundedly rational and have systematic cognitive biases, plus social motivations. This can mean, for example, that a standard policy, such as a Pigouvian tax may not be as effective as regulators expect. Citizens may have difficulty adjusting energy consumption, because of *status quo* bias, or reject it because they feel it is unfair. Experiments can be used to predict these "policy failures" and propose alternatives.

At the same time, the alternatives can also be proposed using behavioral approaches and tested experimentally. For example, social norms have been found to be theoretically relevant and empirically useful to guide behavior. Then experiments have been used to test interventions based on those social norms. People are more willing to make energy savings efforts, or take public transportation, if others are also doing those efforts.

Another area where experiments can help is in the design of novel global institutions to address climate change. A particular case study we have emphasized is that of climate clubs, an original proposal of William Nordhaus that club members commit to lower emissions, and at the same time to have lower tariff barriers towards members than non-members. One important feature of this proposal is its vulnerability to equilibrium multiplicity. Experiments have been used successfully to figure out the factor affecting the probability of different equilibria arising in coordination games. That wealth of information can be brought to bear and help design experiments that address the challenges faced in this specific case.

In the rest of the paper, we present first, in section 2, the ways in which economists have approached the issue of climate change historically, and how the behavioral sciences and the study of social norms have made that approach evolve. In section 3 we review theories and experimental results on how to model social dilemmas and to induce behavior change that have implications for climate mitigation. In section 4 we present the regulatory side, i.e. the strand of institutional design literature, in which we provide a review of different models proposed during the years and our critical standpoint on the existing literature and regarding its future evolution. Section 5 concludes.

2. The standard approach to climate change and the contributions of behavioral science

The proper way to address climate change is that of global public goods: goods whose impacts are indivisibly spread around the entire globe (Faunce, 2012). Clearly, climate is a global public good and climate change is probably the most difficult collective action problem in human history (Barrett, 2018). Collective action problems are also known as social dilemmas[1] and specifically, when the definition is applied to common pool resources, as the climate, they are presented as the "Tragedy of the Commons" (Hardin, 1968). Climate change is a social dilemma in which people must choose between their short-term own self-interest and the longer-term interest of the entire population, operating at multiple scales (individual, national, international).

As a global social dilemma, fighting climate change requires global cooperation at both international and domestic level. At the international level, countries are involved in international agreements which, after rounds of negotiation, bind their actions through mitigation pledges, such as the nationally determined contributions in the Paris Agreement. Yet, the large number and diversity of stakeholders involved in those negotiations makes it difficult to come to an agreement on a policy proposal defining the pledges.

[1] According to Olson (1965), a social dilemma is a situation in which actions that are individually rational can lead to outcomes that are collectively irrational. In other words, is a situation in which individuals would be better off cooperating but fail to do so because of conflicting.

At the national level, each government will need to transform those pledges into a mitigation policy whose success will depend on individuals' behavior and their willingness to cooperate. In order to achieve citizen cooperation in climate commons, policymakers need to understand the reasons that will induce individuals to change their choice, behaviors and lifestyles.

In this framework, it turns out that climate change mitigation can be promoted by resorting to two basic mechanisms: working on social preferences or, in other words, changing players' behaviors and motivations, and creating institutions, by modifying the rules of the game or introducing new ones. In what follows, we give examples of each one of them.

Economists tend to start discussions about climate change and other environmental problems by pointing at public policy (institutional) interventions. These are the systems of laws, regulatory measures, and other actions that governments and parliaments design and implement in response to the problems faced by their citizens seeking to improve their well-being. In a certain way, the objective of public policies can be defined as seeking to limit the social suboptimality of decisions taken "optimally" at the individual level and correct their effects. However, the success of public policies crucially depends on the response of citizens to them. That is, the acceptance, understanding and degree of compliance that citizens make of the measures applied to them. However, there is ample evidence that human beings do not always optimize and, therefore, our decisions often do not result in our own benefit or that of the society to which we belong.

To be concrete, let us start by providing an example. The typical introductory economics textbook reaction to the climate change problem is to suggest the imposition of a "Pigouvian" tax.[2] Such tax increases the price of unclean externality-creating energy sources. If correctly set, at a rate that equalizes the social and individual marginal cost, it would eliminate the social distortion, by aligning the individual and social optimal consumption levels. However, is it really the case that environmental taxes work? A recent meta-analysis by Zhou et al. (2018) shows that "residential electricity demand is almost price-inelastic and income-inelastic in the short-term." This means that a well-meaning environmental tax will not reduce consumption immediately. It will create revenue for the government, and maybe even increase inequality, since energy is a larger fraction of expenditure in poorer households. Does that mean that Pigouvian prices and other standard public policy approaches should be abandoned? Certainly not; following with the same example, Cialani and Mortazavi (2018) show that industrial electricity demand is elastic to price changes. And even Zhou et al. (2018) show that long-term price elasticity is significant.

The previous discussion highlights that society should take a three-pronged approach to address the challenge of climate change. First, we need to intervene to

[2] This is true even at modern/enlightened textbooks such as The Economy by the Core Project. See its unit 12: https://www.core-econ.org/the-economy/

correct wedges between individual and social interests. Second, we need to tackle the individually suboptimal decisions that arise out of a lack of cognitive resources to address a very complex world. And, finally, we need to figure out creative ways to address the global commons problem, which makes climate change so intractable. Since the first path is the common one in Economics, we will concentrate our discussion in this paper in the other two, namely behavior change, and novel proposals to address global commons.

In the last fifty years, behavioral social sciences have successfully shown that many of the actions of human beings are not always governed by the principle of rational optimization. In the pioneering work of Kahneman and Tversky (1974), they show that our decision-making process is often based on automatic decisions ("biases" or "heuristics"), influenced in turn by our emotions, our inability to process information, our behavior in the face of uncertainty or simply because we do not know exactly what is best for us, either in the short or long-term. Behavioral science has therefore based part of its success on taking as its starting point what individuals actually do, rather than a "theoretical" idealization of how they behave (or should behave).

The application of behavioral economics to climate change policy has been considered extensively. Carlsson and Johansson-Stenman (2012), for example, note that the departures from standard ways of thinking take three main forms:

1. Behavior is not motivated just by own material payoffs. Distributional concerns are an important driver of behavior (see Cabrales and Ponti, 2015).
2. Human act in a social context: social approval, norms and status are key motivators (Cole et al., 1992).
3. People have cognitive limitations and therefore sometimes make seemingly irrational decisions. (Simon, 1955).

These considerations suggest margins of action that we will explore in what follows. Thus, Cai et al. (2010) and Svenningsen and Thorsen (2020) have explored the impact of distributional concerns, between and within generations, on the willingness to pay for abatement. It must be remembered at this point that the negative reaction from a part of the French population to a green tax (the *gilet jaunes* movement) was heavily influenced by a perception that the consequences of the tax were not evenly distributed.

Social norms have also been used extensively to modify consumption decisions. Alcott (2011) reports on the OPOWER experiment, where a consulting company, working in conjunction with power utilities in the US used the power of social norms to decrease energy consumption. They sent its letters called "Household Energy Reports" (HERs) with a "social comparison module" and a normative indication as to whether the home is "very good," "good" or "below average." They obtained significant consumption reductions over time. Importantly, they were maintained over time, which indicates that the informational intervention served to assuage cognitive limitations as well.

The final lever of interest is that of innovative proposals to mitigate emissions. The global commons suffer from the standard problem of a social dilemma. If individuals observe that others do not contribute to a public good (and some will surely not contribute) then global contributions will decline over time. There is significant evidence about this for multiple societies (Herrmann et al., 2008).

In principle, the fact that the climate game is repeated can theoretically provide a solution, using the threat of "bad equilibrium reversion" to get cooperation. This is an intriguing, decentralized solution. The trouble is that we have not seen any evidence of it happening in the half century since climate change has become evident; and there are good reasons for this. In a context with many countries, the efficient theoretical equilibrium is too vulnerable to partial agreements between several countries. In a very precise sense, it is not stable against deviations from "coalitions". Perhaps more importantly, from the point of view of human behavior, punishments are difficult to implement because they not only hurt the punished, but also the punisher; and when "this hurts me more than you" it is easy for a group to find a less painful way out.

An obvious solution is to have international agreements where countries commit to reduce emissions. One problem with this approach is its vulnerability to "carbon leakage", the displacement of production and emissions from states with more to less stringent climate policy commitments. One way to deal with this problem is carbon adjustment taxes. All goods entering a trading zone (say, the European Union) will need to pay a tax that is equivalent to the local cost of carbon abatement inside the zone. This neutralizes the perverse incentives to displace production to less regulated areas. This could face limitations because of international trade agreements, but Mehling et al. (2019) show they could be addressed with a careful design. Winchester et al. (2011) show that although the direct effect of carbon border adjustments may be small (and they are a costly way to implement leakage reductions), they can be a useful coercion strategy for countries to implement policies that reduce emissions more cheaply. An evolution of this tool is the one called "climate clubs" proposed in Nordhaus (2015), about which we discuss at length in this paper.

Another potential set of abatement alternatives works through a deeper understanding of the inter-relationship between fiscal and financial climate policies. The financial system's network structure means that climate-induced shocks can spread quickly, but also suggest that the impact of financial policies to fight climate change may spread with the same speed across agents and boundaries, bolstered by the evolution of social norms. The tools of the existing framework (economic, social, legal) for financial stability are sufficiently powerful to change the incentives for firms and reduce significantly their carbon footprints. In upcoming work, Cabrales and Gottardi (2022) explore the impact of misaligned incentives on financial network formation and on social welfare, and Ramos et al. (2022) explore the legal feasibility and opportunity of financial regulatory tools.

3. Experiments on individual behavior

The IPCC, in its Special Report on 1.5 degrees, assumes "behavioral and lifestyle changes" as a vital climate change mitigation strategy complementary to technological measures. According to the IPCC ambitious changes, like how we use fuel, land and other natural resources (Intergovernmental Panel on Climate Change, 2018, 2019) are a priority.

Such an ambitious plan will require a large-scale change that needs to incorporate both the demand (consumers) as well as the supply (firms and politics) side of the society (Alló & Loureiro, 2014). Citizens' responses do not come with negligible resistance. Even policy solutions that focus on technology or structural changes often require behavioral components to succeed, such as the adoption of new technologies or participation in energy-efficiency programs. Ultimately, people must adhere to policies for them to succeed.

Hence the question is: why people fail to engage in behaviors necessary to mitigate climate change? The answer is related to the well-known gap between environmental attitudes and actions that enable a change of behavior. In the literature, we find that the gap between attitudes and behaviors depends on both structural barriers and decision-making barriers. Structural barriers, such as poverty, ill-suited public transportation and services or climate-averse infrastructure, may be lowered with social programs and infrastructure improvements. According to Drew and Van den Bergh (2016), who review the sociopsychological factors underlying the (un)popularity of carbon taxes, the barriers of individuals' decision-making display at both individual level, such as lack of knowledge, ill-oriented motivation, wrong perception about climate change and the effectiveness of mitigation policy, self-efficacy of agents, as well as at the social level, such as generalized trust and social norms and expectations and lack of enabling institutions. Finally, decision making is highly heterogeneous and as such the literature, usually, presents the two barriers as separate, but there exist spillovers effects from structural barriers to psychological barriers.

The scope of this section is to shed light on the barriers that hamper individuals' decision making to enact the behavior change that the climate change mitigation policies address. In the rest of the section, we present how behavior change has been addressed in the literature and the underlying assumptions of individual's decision making which justify the use of top-down or bottom-up interventions. Then, we examine how social norms impact people's ability to cooperate with climate change mitigation. Finally, capitalizing on the previous sections, we present experiments concerning many social dilemma situations that have implications for climate mitigation.

3.1. Behavior change in economic models

If society decides that it has an interest in changing people's behaviors away from unsustainable ones, economic theory has identified four basic options to change

behavior: (1) impose regulations that constrain individual's actions (strategy spaces, in the language of game theory); (2) tax (or subsidize) the undesired (desired) behavior (changing the payoffs, in the language of game theory); (3) other forms of incentivizing (changing the payoffs of) the desired behavior; or (4) provide information.

Neoclassical economists confined themselves to the study of beliefs and incentives, strictly assuming that decision makers are fully rational, self-interested, benefits maximizing, and costs minimizing individuals with stable preferences. Modern behavioral economists relax the homo economicus assumptions when investigating beliefs and incentives and study adaptation of preferences and the role of boundedly rational learning rules.

Given these premises, the approaches to behavioral change of these two literatures need not be viewed as antipodes, but they should reveal the need for effective integration as the IPCC Sixth Assessment Report (IPCC, 2019) suggests and recommends.

We revise the policy mitigation strategies considering both neoclassical and behavioral theory in order to highlight such integration.

1) The first approach to behavioral change consists in drafting regulations, such as bans and restrictions to the use of the harmful alternative. Regulations, such as bans or restrictions, can be implemented more easily and equitably than taxes or incentives. Yet regulations face the problem of needing to be enforced. That is, placing a regulation on an activity does not change the basic internal calculus concerning the individual's actions. If individuals prefer private transportations over public ones, for example, then a law that bans the sale of cars is not likely to be effective unless it appeals to another value (e.g., being a good member of society, adhering to social norms) or there is some enforcement mechanism to ensure adherence.

2) A second approach suggests the use of taxes and subsidies. The rationale behind the use of taxes as instrument for behavioral change rests on the standard economic remedy for internalizing external costs, which is a Pigouvian tax on the pollutant. In this case, what is called for is a carbon tax, levied on carbon-based fossil fuels in proportion to the amount of carbon associated with their production and use. Will such a tax amounts affect people's driving or home heating habits very much, or impact industry's use of fuels? This depends on the elasticity of demand for these fuels and on its regressive effect.

Along with purely economic reasons, interventions such as carbon taxes are shadowed by perils like the lack of public support. Carattini et al. (2018) surveys the literature on qualitative and experimental work highlighting several factors affecting public support. Concerns over carbon taxes stem from overestimation (underestimate) of the costs (benefits) of the tax (Alberini et al., 2018; Heres et al., 2017; Carattini et al., 2018; Odeck & Bråthen, 2002), to equity and pro-social preferences (Bristow et al., 2010; Brännlund & Persson, 2012; Gevrek &

Uyduranoglu, 2015) and to concerns over the implementation and efficacy of the policy (Klok et al., 2006; Steg et al., 2006; Baranzini & Carattini, 2017; Carattini et al., 2017; Hsu et al., 2008; Kallbekken & Aasen, 2010; Kallbekken & Sælen, 2011). In many of the studies above the respondents consider low-carbon subsidies to be a more powerful way to reduce greenhouse gas emissions, especially if the cost of switching from consuming high-carbon goods to low-carbon goods is considered high and given that price elasticity of demand for carbon-intensive goods was considered close to zero. The support for subsidies against the tax can be explained by the loss aversion hypothesis.

A subsidy is environmentally motivated if it reduces directly or indirectly the use of something that has a proven, specific negative impact on the environment. It can take many forms: Value-Added Taxes exemptions on electric cars, feed-in tariffs on renewable energy generation, tax credits for environmentally relevant investment, or provision of public funds for nature conservation projects. Yet all these instruments are not seen as appealing by everyone. The same issues of behavioral biases (Allcott & Rogers, 2014) and bounded rationality (Pollitt & Shaorshadze, 2013) in consumer decision-making prevail.

3) A third approach to behavior change involves providing financial incentives or disincentives for engaging in the desired behavior that are not taxes or subsidies. There are several challenges with using incentives to spur behavioral change.

People's reference points matter in considering a monetary incentive as binding. Whether using incentives as a reward for positive behavior or as a penalty for bad behavior it may be useful to consider loss aversion (the phenomenon that individuals prefer to avoid losses than acquiring gains). It is likely that sanctions, such as fines, are likely to be more effective when framed as losses. On the other hand, ambitious climate-protection goals would require new investments (physical and intellectual) in climate-friendly technologies. These investments are essentially irreversible and people might suffer from the sunk cost fallacy.

Providing monetary incentives raises the total amount that individuals can spend and thus could lead to an increase of consumption of both climate friendly and polluting goods. A further drawback of incentive is their crowding-out effect. It is claimed that monetary compensation can lead to feelings that an activity is not worthy by itself ("intrinsic" motivations) being "crowded out" or partially destroyed. Another problem related to incentives is that the choice over a sustainable behavior might have not only an economic attribute but it could enclose social or symbolic values.

In this regard, individuals are endowed by what Manski (2000) called "preference interactions". Individuals' preference ordering over the alternatives in a choice set depends on the actions chosen by other agents. Such everyday ideas as conformism, jealousy, and paternalism suggest forms of preference interaction. Preference orderings expressed in terms of symbolic values entail attitude, beliefs and identity.

Values influence behavior indirectly by activating norms, thereby creating a feeling of moral obligation to act pro-environmentally (Nordlund & Garvill, 2003). Values generally influence pro-environmental behavior through three different channels: by the perceived importance and likelihood of behavioral consequences, by norm activation and via environmental self-identity (Van der Werff et al., 2014).

4) The fourth approach is based on informational interventions. In reality, all individuals are not endowed with access to the same amount of information or have the same capacity to manipulate the same piece of information: what will be the tool to coordinate them to a new sustainable alternative? In this case, provision of information should cause the individual to realize that the values of the attributes of the promoted sustainable (brown) alternative is better (worse) than they initially believed. Yet lasting behavior change can only occur when informational campaigns provide new information that produces lasting changes in how people view the attributes. There are several impairments to the re-evaluation of the attributes.

The first limitation, given that climate change is a complicated phenomenon, even when people understand and endorse the goal of mitigation policies they may not know which of the many personal changes will be more effective. Attari et al. (2010) shows that within a familiar context like home energy conservation, people often do not know which strategy, between curtailment (e.g., turning off lights, driving less) and efficiency improvements (e.g., installing more efficient light bulbs and appliances), was more *consequential* to achieving energy efficiency. Moreover, the efficient use of energy is a cognitively challenging task and involves trading off short-term versus long-term benefits and costs, and it is unlikely that all consumers can perform the necessary calculations and to arrive at an individually optimal result.

The second obstacle is that even if people know how to act effectively, they may not be motivated to do so. Motivations are linked to how much the climate change problem, with its related costs or losses, is perceived immediate and concrete in both time and space. In order to fight global warming, we ask individuals to make an immediate costly effort to regulate their behavior today (e.g., to drive less, to consume local, to use an alternative energy resource) but the rewards from these efforts, e.g., in form of better environmental conditions, are only realized in the future. Such self-regulatory problems are very difficult to address (Weber, 2006), even in a context in which individual's self-interest is clearly at stake (e.g., increasing healthy eating to promote long-term health, Hall and Fong, 2006). A simple informational strategy would not be effective.

The third hurdle is related to the idea that people have rational expectations about the consequences that one action will produce given the informational set available at the time of the decision. Yet, neoclassical economics also assumes that people would form such expectations from observational learning, i.e. from observation of the actions chosen and outcomes experienced by others. Observational learning is highly heterogeneous in reality. If it is true that people learn from what they observe,

people living in the same city but in different residential areas would have a different glimpse of the reality. General information about how society at large adopts more sustainable behaviors might not have the same effect as information based on the behavior of local communities. Any information colliding with such experiential window would be discarded. For example, Allcott and Mullainathan (2010), Allcott (2011), and Allcott and Rogers (2014) show how the power of social comparison in home energy reports (HER) are a cost-effective climate policy intervention to push people towards sustainable energy consumption or providing households with tailored information regarding their energy use reduces their energy use (Abrahams et al., 2007). In contrast, direct repeated information about the causes of climate change does not lead to a reduction of carbon consumption (King et al., 2009).

Hence, the nature and the saliency of the information delivered, and the mode of information delivery is relevant. Chetty et al. (2009) finds that consumer reactions to taxes depends on the visibility and salience of the tax. By the same token, assuming that any type of information will be accounted for in the decision process is highly unrealistic too. In behavioral economics there exists a phenomenon called information avoidance (Golman et al., 2017) which refers to situations in which people choose not to obtain knowledge that is freely available. Active information avoidance includes physical avoidance, inattention, the biased interpretation of information (see also confirmation bias) and even some forms of forgetting. Alcott (2011), for example, indicates that 40% of US consumers do not consider vehicle gasoline when they face the decision to purchase a vehicle.

To date, the literature on behavioral approaches to mitigation policy has focused its attention on how to correct all the previous behaviors labeled as far from the neoclassical assumption, but none on understanding whether individuals' competences are up to the challenge of behavioral change (Kolle, 2015). The literature, so far, looks at the individual as a boundedly rational, boundedly self-interested being who suffers from lack of willpower (Thaler & Sunstein, 2008; Halpern, 2015). Consequently, the policy approach that has been used is the one of nudging such as default option (Allcott & Kessler, 2019; Bernheim et al., 2015).

Psychological barriers are not only based on limited cognition or on social interactions. There are individual psychological barriers that need to be added in the analysis to really assess people's motivation to take part to the climate change battle. The inertia of people to change their behavior emerges, not only as a consequence of increasing returns to conformity but also because of the way in which individual self-views evolve. If most members of the population do not believe in their abilities to change the *status quo* they are more likely to unquestionably follow the standing behavioral rule (Olson, 1965). For instance, perceived lack of control predicts the choice between public transportation and driving (Kaiser & Gutscher, 2003; Heath & Gifford, 2002). Thus, when individuals perceive little control over the problem of climate change, they may fail to act. Specifically, there is evidence that low-income individuals are more likely to respond to curtailment policy and to stick to default choices (Ghesla et al., 2020). In particular, investigating electricity contract choices

Hortacsu et al. (2017) find that households with lower income and lower education are less likely to switch their electricity contract.

In addition, the policy approach changes if we consider individuals only as cognitively and socially bounded actors or if we take into consideration also their competences. In fact, by addressing only individual's cognitive deficiencies, policymakers can steer (nudge) individuals' behavior toward behaviors that are consistent with their ultimate goals or preferences – and that result in better outcomes than would otherwise be obtained (Rebonato, 2012; Thaler & Sunstein, 2008). In doing so, the architect does not aim to foster people's competences for making better choices. Hertwig and Grüne-Yanoff (2017) propose a different type of intervention to complement the nudges, what the authors called boosts. Boost interventions target competences and capabilities rather than immediate behavior. By fostering existing competences or developing new ones, boosts are designed to enable specific behaviors. Furthermore, they have the goal of preserving personal agency and enabling individuals to exercise that agency. Therefore, in order to improve one's judgment boost interventions aim to train and strengthen the cognitive system by promoting, for example, (i) statistical, energy, and financial literacy[3], (ii) deliberative skills, or (iii) the use of evidence-based guidelines. For example, Kalmi et al. (2020) show that energy-related financial literacy might guide consumers' decisions toward energy efficiency and conservation.

3.2. Behavior change and social norms

As recently argued by Fehr and Schurtenberger (2018), a large variety of behavioral regularities with regard to human cooperation can be explained by a significant share of individuals adhering to a social norm of conditional cooperation (see Kimbrough and Vostroknutov, 2016; Kölle et al., 2020; Szekely et al., 2021) for direct evidence on the importance of norm-following for cooperation).

Prior work has shown that harnessing social norms can be instrumental in addressing large-scale social dilemmas (Ostrom, 2000; Bicchieri, 2005; Biel & Thogersen, 2007). People's understanding of collective action problems does not occur in a vacuum because its interpretation is influenced by values and beliefs shared in groups for which they feel a sense of belonging. Hence, the first psychological barrier derives by the presence of social norms. Social norms are generally understood to be shared rules of conduct that are partly sustained by approval and disapproval (Elster, 1989) or ideal form of behavior to which individuals in a social group try to

[3] This competence can be achieved through (a) graphical representations (Lusardi et al., 2017); (b) experienced-based (as opposed to purely description-based) representations (e.g., Kaufmann et al., 2013); (c) representations that avoid biasing framing effects (Spiegelhalter et al., 2011); (d) training in transforming opaque representations (e.g., single-event probabilities) into transparent ones (e.g., frequency-based representations, Sedlmeier and Gigerenzer, 2001); and (e) training of general math skills (e.g., Berkowitz et al., 2015).

conform (Young, 2015; Burke & Young, 2011). They form as the unexpected result of individuals' interactions that, through learning, specify "what is acceptable and what is not in a society or group" (Muldoon et al., 2014). Moreover, social norms are self-enforcing at the group level because people adhere to certain norms of behavior if these norms make them better off, or meet their needs, but also if they expect others to adhere as well.

In a coordination game, a social norm corresponds to a pure equilibrium of the game that is played repeatedly by members of a population with or without a punishment of the deviants. The relevant point is that the equilibrium holds at the population level, inducing common expectations and behaviors for an interaction that is repeated over time by members of a social group. In Manski's words (Manski, 2000), a social norm would be the result of people's interactions colliding in a dynamic of expectation coordination.

At the same time, social norms do not only coordinate people's expectations and, eventually, judgement, but also people's preferences (Manski, 2000; Bicchieri, 2005). In other words, an agent's utility may derive simultaneously from his personal and idiosyncratic preference for a particular action, and from his preference for conformity to the actions of his reference group (Akerlof & Kranton, 2010). Norms and personal preferences are also highly intertwined, as norms can shift motivations; and even in cases where norms do not shift personal beliefs, they still can have a substantial impact on personal behavior (e.g., Paluck, 2009a). Moreover, according to Paluck (2009b), it is often easier to impact perceptions of social norms than directly shift people's attitudes or beliefs on a topic. Consequently, an agent with such preferences would react to incentives that are norm-based, or will follow taxes that are in agreement with the local social norms, or will accept information that are salient with the features of the norms.

We could look at social norms as the solution of global social dilemma. Coordination of expectations is extremely important in global dilemma, such as the climate change, where uncertainty and ambiguity related to the effectiveness of agents' effort is a hurdle to overcome. At the same time, coordination of expectation reduces agents' worries for the free-riding of others and align people's incentives (whatever those are).

In fact, if we analyze the expectations features of the social norms, we soon realize that norms exists if two kinds of expectations are formed in the mind of people. According to Carattini et al. (2020), social norms have different levels of visibility.

Empirical expectations[4], that are the most visible ones, are expectations about what people do. By observing other people cooperating, agents start forming expectations about the fact that cooperation is a widespread behavior in the social reality in which agents live (Allcott & Rogers, 2014). Normative expectations, the less visible ones, are those related with what people should do in respect of the context, the reference group and the decision to take. Normative expectations make people's beliefs

[4] Bicchieri (2006) calls empirical expectations what Cialdini (2003) descriptive norms.

converge towards what is perceived as the socially right behavior to have (Schultz et al., 2007; Szekely et al., 2021). For Bicchieri (2006, 2016) both types of expectations are necessary to see the emergence of social norms.

For all of these features and dynamics, norms are likely to be an apt solution to addressing climate change because they are a robust source of influence (Ostrom, 2000; Bicchieri, 2002; Biel & Thogersen, 2007; Bolsen et al., 2013; Nyborg et al., 2016; Huber et al., 2018). Of course, to be part of the solution to climate change and not its hurdle, new green and sustainable social norms need to be created or helped to emerge within the society. And to do that, many policy instruments are needed.

Examples of how empirical expectations induce behavioral change are the following. In Allcott and Mullainathan (2010), Allcott (2011) and Allcott and Rogers (2014) comparison in home energy reports (HER) works as signal of the behavior of others pushing people towards the creation of empirical expectation on sustainable energy consumption. Baranzini et al. (2017a) find that Swiss consumers are more likely to adopt solar panels if neighbors have already done so. Such diffusion is driven by imitation of conspicuous consumption and communication of positive information. The same effect in the adoption of rooftop photovoltaic technology has been identified in California (Bollinger & Gillingham, 2012), Connecticut (Graziano & Gillingham, 2015) and Germany (Rode & Weber, 2016). All of the above confirm the fact that rooftop solar photovoltaic panels were visible, they convey information about the behavior of others in a given community (i.e., the local social norm). Even though there were also financial incentives driving the adoption, the literature suggest that the recurrent view of the panel worked as a reminder of the widespread social behavior.

Empirical expectations do not suffice, by themselves, to sustain a change of behavior for a long time. Normative expectations are necessary too. In fact, as found in Székely et al. (2021) individuals' cooperative behavior is primarily sustained by both empirical and normative expectations of cooperation, which are formed through individuals' social interaction. Also, Schultz et al. (2007) apply this approach to a field experiment on household energy consumption in California finding that when interventions are based on both empirical expectations and normative ones, people tend to replicate the behavior of the more cooperative individuals in their local context.

Social norms have also been highlighted as a means for overcoming the limit faced by regulations in achieving behavior change. In fact, social norms help in building public support for climate policy (for a review, see Alló and Loureiro, 2014; Sparkman et al., 2021). Furthermore, social norms have been shown to influence one's policy attitudes, even when those norms are contrary to one's initial personal beliefs (Todorov & Mandisodza, 2004). Norms also shape support for climate change policy measures among policy-makers themselves (Nilsson et al., 2004). Beyond policy support, social norms have been shown to impact whether citizens are likely to engage in political action on climate change, such as contacting government officials, voting for "green" candidates and protesting (Doherty & Webler, 2016).

The next section describes how climate change can be modelled with specific experimental setups and how social norms could be used as instruments to steer people's mitigating behavior.

3.3. An experiment on behavior change for climate mitigation

For years the experimental literature on social dilemma has focused on using standard public good games to study how to sustain cooperation. Standard public good games are concerned with the creation of a collective gain (Sturm & Weimann, 2006; Fehr & Gächter, 2000; Fischbacher & Gachter, 2010). Even though climate change is considered an example of a global public good, it has specific characteristics that require a different representation than standard public good games. In fact, climate change is more about avoiding an uncertain public bad, rather than the creation of a collective gain.

The literature about sustaining cooperation for addressing climate change is divided into two strands of literature: the static repeated games and the dynamic social dilemma.

In the set of static repeated games, we find games that model climate change as a collective-risk social dilemma, i.e., a problem of sustaining cooperation when facing an emission threshold that may result in a catastrophe (Milinski, 2008), while others model it with an incremental damage from pollution (Ghidoni et al., 2017).

This "collective-risk social dilemma," is a threshold public good game of loss avoidance played with sequential contributions to a fund aimed at avoiding a probabilistic loss arising if the target is missed (Milinski et al., 2006; Dreber & Nowak, 2008; Chakra et al., 2018).

This model belongs to a larger set of dilemmas also known as threshold public global good (Pacheco et al., 2009). At the start of the game, participants are each given an endowment, and they must decide whether to contribute, up to a predefined amount, to the common good over a fixed number of rounds. If the joint contributions of all the participants over those rounds are equal or above a certain threshold, then the disaster is averted, and they receive as a reward the remainder of the endowment (hence the dilemma). On the contrary, if the target is not reached, there is a probability that a disaster may occur, resulting in an economic loss for all the participants (they lose the remainder of their endowment). In the experiments, people only tend to contribute to avoid the disaster if they perceive the risk to be high (Hagel et al., 2016; Milinski et al., 2008). Moreover, even when the risk is high, theoretical models indicate that players should delay their contributions until the moment when the disaster is known (Abou-Chakra & Traulsen, 2012; Hilbe et al., 2013).

The phenomenon of climate change is very well depicted through the features of the collective-risk social dilemma. The risk parameters, the threshold, and the loss avoidance construct, which make the game non-linear and the collective benefit

uncertain as it is only achievable in the future. For this reason, the game has shed new light on the issue.

Yet, in a real-world scenario both the amount (threshold) as well as the timing when it has to be achieved are uncertain, as they are based on predictions and thus inherently suffer from uncertainties. Prior work on uncertainty about what amount (threshold) should be achieved in such games and more so in case of ambiguity (Barrett & Dannenberg, 2012, 2014; Dannenberg et al., 2015) has shown that the level of cooperation, i.e., the willingness to contribute in both games, is negatively affected. Uncertainty about the timing in which a predetermined target yields benefits decreases cooperation (Jacquet et al., 2013; Kolle & Lauer, 2020).[5]

Moreover, the challenge of this game is coordination. Players are best off when synchronizing contributions in the face of multiple equilibria. The game therefore calls for an instrument able to facilitate such coordination. Some authors use communication as an instrument for coordination. Tavoni et al. (2011) show that income inequality and the ability to communicate also affect the frequency of avoiding a catastrophe: success is more likely in groups making choices that reduce inequality and are able to communicate.

Other authors use social norms as a coordination mechanism. Szekely et al. (2021) design a repeated threshold public goods game with elicitation of social norms and social norms strength showing the causal evidence that social norms change in response to threat variants and that stronger norms increase social coordination. To do so the experiment lasted for 30 days to allow social norms of cooperation to emerge and to be enforced. During the experimental days, the agents' expectations (both empirical and normative) were elicited and agents' social norms strength was computed. Agents were exposed to two different threats where the risk of a catastrophe was either high or low. Social norms associated with the two risk scenarios were the tools driving people cooperation.

The main conclusions of the experiment can be summarized as follows: The authors find that in the scenario where the risk was high, social norms strength was higher, pushing agents to cooperate more. The positive relation between social norm strength and risk answers the daunting question of how to address cooperation in uncertain collective-risk social dilemma. Moreover, such relationship would persist in time: the paper results show that in a high-risk scenario the effect of higher social norms strength would persist also when the risk lowered. Finally, social norms were not only endorsed by agents, via increasing cooperation, but they were also enforced by the latter, leading to punishment actions against the norm-breakers.

[5] These results hold true also when treatment in delayed payment is not present and they hold true in both the lab(oratory) as well as in the field. Fehr and Leibbrandt (2011) find that time preferences measured in the lab predict cooperation behavior outside the lab in a situation that entails an intertemporal component. In particular, they find that more patient fisherman use more sustainable fishing instruments that are less likely to exploit the collectively used fishing grounds. Similar evidence is provided by Boonmanunt et al. (2020) who show that time preferences elicited in the lab predict replenishment behavior in the field.

The above experiment aligns with the studies of dynamic setups that have been carried out recently. Climate change externalities are dynamic because they depend on the stock of pollution accumulated in the atmosphere and not just on the yearly flow. Cooperation in dynamic set-ups appears more difficult than in static ones. The two main references in this area are Battaglini et al. (2016) and Calzolari et al. (2018).

As Battaglini et al. (2016) point out, there are two main differences in dynamic public good problems with respect to static ones. First, there is now a "dynamic free-rider" problem, where an increase in the contribution of one agent in the present triggers a decrease in the contribution of others in the future. The other problem is that there may be a large number of dynamic equilibria. The paper offers a number of important methodological contributions. However, from our point of view, the main result is on the equilibrium selection. The "good" equilibria where player use strategies that react to actions other than the accumulated level of the public good is not observed in the data. Instead, the Markov perfect equilibrium (usually Pareto inefficient) is the typical observed outcome.

Calzolari et al. (2018) uses a more specific dynamic public good game that mimics climate change more closely. Their main observation for our purposes relates to a situation where conditions are close to the ones in reality: the stocks of the harmful action are long-lasting. In this case they observe that participants in the experiment cooperate strongly at the beginning of the game, but then it decreases in a very significant way.

In summary, the insights obtained from the dynamic models corroborate those already found in the static games, to a large extent, but emphasize just how difficult is to obtain cooperation in dynamic situations.

4. Institution design

4.1. The role of institutions in climate change

Since the last quarter of the 20th century, there has been quite some research on the theory and modeling of global public goods that, as we have discussed above, is the way climate change should be addressed. Among this large body of literature, quite a few papers discuss the game theory and modeling of coalitions (e.g., of institutions) of countries. While a thorough review of this literature is beyond the scope of this paper, we here summarize the most important results along this line to provide a proper context for our focus paper, namely Nordhaus (2015). A general survey on treaties about global public goods preceding Nordhaus work can be found in Barrett (2003).

Prior to Nordhaus' work, a first relevant contribution was that of Carraro and Siniscalco (1993), who analyzed the problem of free-riding within international agreements for global public goods. Their key finding was that only a small number of countries would take part in the agreements as they were designed, and only if it was possible to make and enforce binding commitments. This result would arise

also from many other studies, leading to Nordhaus calling it the "small coalition paradox". Subsequently, Chandler and Tulkens (1995) and Chandler (2007) showed that transfers between participants are needed in general to have stable cooperative equilibria, although in a few special cases such transfers might not be necessary. However, they assumed that any single defection is enough to break the coalition, something like a doomsday scenario, to prevent defections from participation. The drawback is that this strategy works against the punisher as well as the punished, and hence the agreement is not very appealing (technically, they are not renegotiation-proof). Along these lines, Yang (1999, 2008) considered how transfers could improve the overall abatement, finding that it requires substantial transfers from North to South to induce cooperation.

Bosetti et al. (2012) studied the problem using the WITCH (World Induced Technical Change Hybrid) integrated-assessment model (Bosetti et al., 2006), a global dynamic model integrating the interactions between the economy, the technological options, and climate change (it is worth mentioning that WITCH is an open-source model available for any further studies and still running today, see https://www.witchmodel.org). Their main finding is that only a global coalition of all regions is able to control the amount of greenhouse gas in the atmosphere, but unfortunately, such a global coalition turns out to be unstable even with monetary transfers. The small coalition paradox showed up again as they also found that smaller coalitions can be stable but cannot lead to efficient climate change mitigation.

Related results have been found in Finus et al. (2005), who showed that no non-trivial coalition is stable if membership is open, and by Weikard et al. (2009) who also noticed stability problems that cannot be prevented by transfers, and that bargaining over them can generate coalition instability.

In 2015, Nordhaus' paper was published, proposing a mechanism to allow this kind of institutions to work without resorting to unrealistic assumptions, which we discuss in detail below. A lot of discussion ensued, leading to recent commentary papers such as Tagliapietra and Wolff (2021), where it is claimed that if the three biggest economies would agree on a carbon tax on imports, their agreement would catalyze global climate.

As for more academically oriented literature, several papers are particularly relevant. Thus, Vogt (2016) considered a situation with heterogeneous actors that at the same time are inequality averse. They applied their results, estimating empirically some of their model parameters, to the problem of climate mitigation policies using the twelve world regions from Nordhaus' RICE (Regional Integrated Climate-Economy) model. Their conclusion aligned with the instability problems of coalitions, as they found that wealthy countries have economic incentives to leave a coalition, even taking into account a preference for advantageous inequality, whereas poor countries also improve their welfare leaving, as both their absolute payoff increases and their disutility from disadvantageous inequality is reduced. As in previous research, suitable transfer schemes can stabilize coalitions formed by economically divergent members.

Another issue that has been considered in the literature is the fact that it is often the case that belonging to one or another group is not a completely voluntary choice. In this context, Dannenberg and Barrett (2018) showed that learning within a group takes place slowly in time, hindering the emergence of cooperation. Indeed, groups can fail to implement an efficient institution because of expectations that it would not work, or, if the institution is implemented, when not enough members realize its advantage, the institution eventually breaks down. This leads to further pessimism and causes groups to accept their fate. Heitzig and Kornek (2018) studied a related situation, in which countries expect that if they do not take part in a coalition, others might, showing that this implied once again poor prospects for collaboration. However, they also show that in a dynamic setting, an efficient coalition is achieved when players are sufficiently far-sighted or there is an immediate coordination caps right after market linkage.

More recently, Nordhaus himself has revisited the problem, including now a repeated game perspective. In Nordhaus (2021), the author extends the one-shot approach to many periods, introducing an approach that deals with "supportable policies" in a scenario of multiperiod clubs. An additional novelty of his study is that he considers interaction between club effectiveness and rapid technological change, neither of which will allow to attain the objectives of international climate policy on its own. Trade sanctions will be too costly to produce deep abatement in the absence of accompanying, rapid technological innovation, while innovation alone is still subject to countries free riding. Interestingly, he shows that when the two factors work together, international climate goals can be achieved.

In addition, Karatayev et al. (2021) address the problems of large-scale negotiations, when commitment to mitigation is costly and uncommon, and demonstrate that a well-timed policy shift from local to global legally binding agreements is much more effective than using only local, only global, or both agreement types simultaneously. The reason is that local agreements foster commitment and mitigation in early adopting groups, and subsequently global agreements bring in late-adopting groups.

As can be seen from the above, necessarily brief, summary, the idea of coalitions and similar institutions to fight climate change, in which Nordhaus proposal of climate clubs is framed, has been the subject of much theoretical research. However, the question naturally arises as to the real applicability of all those results, particularly because more often than not, results are negative (meaning that coalitions either do not form or are unstable) and because many mechanisms proposed to deal with these issues have not been tested at all. Here is where experiments are needed and are, in fact, the only way to bring these proposals closer to implementation. In the next section we discuss how experiments are actually informing the alternatives available for climate change abatement policies.

4.2. Some background on experimental approaches to institutions

One of the first questions that was experimentally studied in this context was the effect of inequity aversion. In McEvoy and Stranlund (2016) the problem of coalition-forming in the presence of players that are averse to payoff inequality between coalition members and outsiders was considered. Their laboratory experiments showed that the bigger the gap in payoffs, the less likely coalitions between members and freeriding non-members are. Importantly, they designed their experiment in order to prevent confounding effects arising from the interplay of inequality and the smallest size for profitable coalitions. The main conclusion is then that controlling for the participation threshold size, coalition formation rates decay when the payoff gap between members and non-members increases, making it difficult for the coalition to hold for long.

Bosetti et al. (2017) addressed the issue of effort coordination between coalition members and non-members when they are threatened by a catastrophe. Agents interested in having the coalition formed may commit some of their investments to a climate change related project that offers smaller payoffs as a signal of their commitment. In their experimental design, externalities cannot be totally internalized by the countries joining the coalition, and second-movers' contributions are needed to avoid catastrophic losses. By modifying the returns of the two investments and the diffusion of the gains to second movers, they found that a sizeable coalition of early investors in the clean technology is more likely if benefits are appropriated by the members, and that in fact spillovers can bring in second-movers.

Schmidt and Ockenfels (2021) focused on a proposal by Weitzman (2014) to change the negotiation focus to a uniform common commitment (e.g., a minimum price for carbon) that would promote international cooperation. In their experiment, human subjects participate (voluntarily) in a public goods game and differ in benefits and costs. Irrespective of treaties being enforceable or self-enforcing, it turns out that negotiating a uniform common commitment is better than negotiating individual commitments (as in the Paris Agreement) to promote cooperation, and it is also better than commitments tailored to specific situation of each party (as was the case with the Kyoto Protocol). Finally, another very recent paper (Dong et al., 2021) analyzed the impact of a financial incentive for developing countries to reduce carbon emissions. They observe that such financial incentives lead to higher global contributions towards emissions reduction and effectively reduce emissions even without binding enforcement. This suggests that developed countries should devote some of their resources and incentivize developing countries to reduce their emissions.

These experimental results indicate that, indeed, tests of theoretical proposals to mitigate climate change through institutions can be done while, at the same time, indicate that experiments have to be designed very carefully if their conclusions are to shed light on the applicability of the proposed institutions. It is then worth to discuss a specific example in detail in order to better understand how the interplay between

theory and experiments may be most fruitful, and this we do now by considering the case of Nordhaus' climate clubs.

4.3. Nordhaus' proposal of climate clubs

Nordhaus' (2015) proposal arises from the realization that very many international conflicts have been solved through international agreements, even if international law, as arises from the 1648 Treaty of Westphalia, is an important obstacle. Indeed, the key feature of international law is that countries are all legally equal, and that they must join international treaties voluntarily. For the case of climate change, the temptation to free-ride on other countries' efforts to mitigate it is very clear, and there is little incentive in principle to join a treaty dealing with this issue.

To solve this dilemma, Nordhaus resorts to the theory of clubs (see Sandler and Tschirhart, 1980 for a review). According to this theory, clubs are voluntary groups that yield benefits for their members when there is a cost of producing a public-good-type resource that can be shared. For such an institution to be successful, it is necessary that the corresponding arrangement, including the dues, benefits all its members. In addition, non-members can be sanctioned and that these sanctions inflict a relatively low cost to members. Of course, for the club to work membership should also be stable, so members do not leave it. Examples of actually existing such clubs are international free-trade treaties or military alliances. In these two cases, there are costs, such as low trade barriers or the cost of sustaining an army and defending other members. Nordhaus's proposal is directly inspired by these examples. Still, he presents it as an idealized solution that will never exist in its pure form; the hope is that it opens a way to come up with a system that overcomes free-riding in the context of climate change mitigation.

The key idea behind the climate club is that members agree to undertake harmonized emissions reductions. For example, countries belonging to the club may commit to implement policies leading to a minimum domestic carbon emission price ($25 per ton of carbon dioxide in Nordhaus, 2015). The figure itself is not relevant for the discussion, but nowadays it is of course outdated and prices in the European Union at the time of writing are approaching $100 per ton of carbon dioxide. Countries would be at liberty to choose their own specific mechanisms (such as carbon tax, cap-and-trade, or hybrid designs).

Crucially, countries that are not members of the club are penalized. Nordhaus suggested that members could impose uniform percentage tariffs on the imports from non-members into the club territory. In this manner, a strategic situation is created so that countries acting in their self-interest will join the club assuming very high levels of emissions reductions. The advantage of Nordhaus' system is that countries that do not comply can expect a punishment that is credible, because the members of the club benefit from it. In his paper, Nordhaus examined in detail the club and then considered an empirical model to show that it could actually work. His simulated

model, with parameters taken from climate models, allows him to investigate the characteristics of stable climate clubs for different emission reduction targets. The main result he obtains from the model is that significant reductions in emissions can be obtained with clubs that are stable.

When looking at the climate club design in detail, it turns out that one potential problem with it is that, in principle, multiple equilibria are possible. Indeed, if a country anticipates that no other country will join the club (or only a few, not really significant ones that can affect it with their tariffs), it has no interest in joining. Therefore, it is important to check how severe this problem can be, and here is where experiment can bring a substantial contribution to this literature, by studying whether or not clubs do form. The next section describes a proposal for an experimental design that could address this question.

4.4 An experiment on Nordhaus' climate clubs

Our proposed design incorporates as much as possible the characteristics of the climate club idea. We believe that the design should contain at least two important aspects. One is the fact that in reality countries are heterogeneous. Some are bigger and/or wealthier than others are, and thus have a different impact on climate and on trade. The other is that Nordhaus did not specify the process by which the clubs form. There are many possibilities, and we know from past literature than the protocols for coalition formation are crucial to determine which coalitions actually form (see e.g. Rogna, 2019).

In the experimental setup we are proposing, participants play a number of rounds within a group always composed by the same six people. Within each group two participants are informed that their profile is A, while the other four participants have profile B (and they know it). The *profile* is a neutral word to convey the heterogeneity we mentioned before. The A-profile participants are the "wealthy/large" ones and will have a larger endowment than the B profile players will.

As mentioned, we believe the group formation is a critical part of the experiment. We describe now the general structure of the process and the different experimental treatments. Every round consists of several stages, beginning with a *pledging stage*, in which participants decide whether they want or not to make the pledge of joining the club. In fact, this design is the explicit form of what we believe is a process by which countries sign "expressions of interest" and then enter the club after observing others, as a cheap and effective way to implement the club. This pledge decision will remain in their history of pledge decisions over the entire experiment, and participants have the history of pledge decisions of groupmates available during the entire experiment. The same will occur with the rest of decisions participant have to make in a round. We note that at this stage, participants that make the pledge they are not yet committed to join the club during this round.

The next stage is the *implementation,* played only by participants who pledged to join the club in the previous stage, while the rest of the participants simply skip this stage. During the implementation stage, participants decide whether they want to ratify the pledge of joining the club, having information on the number and type of others who have made a positive pledge. This stage is crucial, and it can possibly take different forms. The club could be formed only if there is unanimous confirmation by pledgers. Alternatively, it can form only with those who confirm.

Finally, there is the *contribution stage*. If the club is formed in the previous stage, all participants within the club are forced to contribute their entire endowment to the common pool. If the club is not formed, they have to decide how many points of their endowment they contribute. Participants who did not join the club from the start also make their decisions to contribute at this point. Subsequently, the amount of the pool is multiplied by some factor and shared in proportion to the participant type/size; participants are informed of everybody's earnings and the next round begins.

To unveil the effects of the different elements of the club formation mechanism, we should consider four different treatments of the experimental design we have just summarized. Thus, in the baseline treatment T1, there is only the contribution stage, i.e., it is a (heterogeneous) public goods game, without any club. Treatment T2 allows clubs to be formed even if some of the members who initially pledged to join withdraw their pledge, whereas in treatment T3 clubs are only formed by unanimous ratification of all pledges. Both in T2 and T3 there are benefits to club members but no punishment to non-members. Treatment T4 introduces the punishment to non-members, and at the same time it keeps the unanimity rule. We note that this last treatment is the one that reflects closer the spirit of Nordhaus' proposal, so we will refer to it as the "climate club" treatment.

We have run the experiments and done some preliminary analysis of the results. There are several tentative conclusions that emerge from that analysis. One is all the institutions (T2, T3 and T4) deliver improvements in contributions with respect to the baseline T1. But, importantly, the full clubs treatment T4, which includes sanctioning, stabilises cooperation levels and make them resilient. It also provides strong incentives to form large clubs of highly committed participants.

A second important observation is that the effect of sanctions is more evident on highly endowed (wealthy) participants, which may suffer from the punishment when all the poorer subjects form clubs, leaving them alone. This hints at the idea that clubs could be a tool for poor countries to put pressure on rich ones. However, this conclusion should be taken with caution. The economic and power imbalance between developed and less developed nations in real world is very different than in our stylized version.

Finally, while high levels of cooperation can be reached without the rule of unanimity for clubs to form, when this rule is present the clubs are more robust and initial pledges to participate are almost always honored.

5. Conclusions

Cabrales et al. (2022) find an embarrassing disinterest in climate change by academic economists (measured by the total number of articles published in the so call "top-5" economic journals). This fact can be explained by the existence of a clear and well-understood set of policies. Namely, we have Pigouvian taxes and subsidies, regulations, and, if push comes to shove, markets for permits. If you want redistribution to compensate for the effects of those taxes, you can always do it through the tax and welfare system.

Perhaps the most important insight we can gather with this review is that policymakers and citizens have a useful tool to complement the existing ones in the pursuit of "safe and effective" policies for what is perhaps the defining problem of our age. We have shown first that there are good reasons why societies may oppose the standard tools economists want to use to address this problem. But, more hopefully, we show that there are other tools that may work as well. We have shown we can exploit the influence of the community through social norms. We can use behavioral interventions that promote the desirability of policies. We can also leverage the fact that climate shocks can have systemic financial consequences and use the mighty arm of financial regulation.

The other important insight is methodological. Theory and numerical simulations are of course complements and need to be used side by side with experiments, but they are not enough to get a good understanding of the challenges and threats we face in the future.

References

Abou Chakra, M., Bumann, S., Schenk, H., Oschlies, A., & Traulsen, A. (2018). Immediate action is the best strategy when facing uncertain climate change. *Nature Communications*, *9*(1), 1-9.

Abrahamse, W., Steg, L., Vlek, C., & Rothengatter, T. (2007). The effect of tailored information, goal setting, and tailored feedback on household energy use, energy-related behaviors, and behavioral antecedents. *Journal of Environmental Psychology*, *27*(4), 265-276.

Akerlof, G. A., & Kranton, R. E. (2010). *Identity economics*. Princeton University Press.

Alberini, A., Ščasný, M., & Bigano, A. (2018). Policy-v. individual heterogeneity in the benefits of climate change mitigation: Evidence from a stated-preference survey. *Energy Policy*, *121*, 565-575.

Allcott, H., & Mullainathan, S. (2010). Behavior and energy policy. *Science*, *327*(5970), 1204-1205.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, *95*(9-10), 1082-1095.

Allcott, H., & Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, *104*(10), 3003-37.

Allcott, H., & Kessler, J. B. (2019). The welfare effects of nudges: A case study of energy use social comparisons. *American Economic Journal: Applied Economics*, *11*(1), 236-76.

Alló, M., & Loureiro, M.L. (2014). The role of social norms on preferences towards climate change policies: A meta-analysis. *Energy Policy*, *73*, 563-574.

Attari, S. Z., DeKay, M. L., Davidson, C. I., & De Bruin, W. B. (2010). Public perceptions of energy consumption and savings. *Proceedings of the National Academy of Sciences*, *107*(37), 16054-16059.

Baranzini, A., & Carattini, S. (2017). Effectiveness, earmarking and labeling: testing the acceptability of carbon taxes with survey data. *Environmental Economics and Policy Studies*, *19*(1), 197-227.

Baranzini, A., Carattini, S., & Péclat, M. (2017). *What drives social contagion in the adoption of solar photovoltaic technology* (No. 270). Grantham Research Institute on Climate Change and the Environment.

Barrett, S. (2003). *Environment and statecraft: The strategy of environmental treaty-making: The strategy of environmental treaty-making*. OUP Oxford.

Barrett, S., & Dannenberg, A. (2012). Climate negotiations under scientific uncertainty. *Proceedings of the National Academy of Sciences*, *109*(43), 17372-17376.

Barrett, S., & Dannenberg, A. (2014). Sensitivity of collective action to uncertainty about climate tipping points. *Nature Climate Change*, *4*(1), 36-39.

Barrett, S. (2018). Choices in the climate commons. *Science*, *362*(6420), 1217.

Battaglini, M., Nunnari, S., & Palfrey, T. R. (2016). The dynamic free rider problem: A laboratory study. *American Economic Journal: Microeconomics*, *8*(4), 268-308.

Berkowitz, T., Schaeffer, M. W., Maloney, E. A., Peterson, L., Gregor, C., Levine, S. C., & Beilock, S. L. (2015). Math at home adds up to achievement in school. *Science*, *350*(6257), 196-198.

Bernheim, B. D., Fradkin, A., & Popov, I. (2015). The welfare economics of default options in 401(k) plans. *American Economic Review*, *105*(9), 2798-2837.

Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.

Bicchieri, C. (2016). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.

Biel, A., & Thøgersen, J. (2007). Activation of social norms in social dilemmas: A review of the evidence and reflections on the implications for environmental behaviour. *Journal of Economic Psychology*, *28*(1), 93-112.

Bollinger, B., & Gillingham, K. (2012). Peer effects in the diffusion of solar photovoltaic panels. *Marketing Science*, *31*(6), 900-912

Bolsen, T., Leeper, T. J., & Shapiro, M. A. (2014). Doing what others do: Norms, science, and collective action on global warming. *American Politics Research*, *42*(1), 65-89.

Boonmanunt, S., Lauer, T., Rockenbach, B., & Weiss, A. (2020). Field evidence on the role of time preferences in conservation behavior. *Journal of Environmental Economics and Management*, *104*, 102368.

Bosetti, V., Carraro, C., Galeotti, M., Massetti, E., & Tavoni, M. (2006). A world induced technical change hybrid model. *The Energy Journal*, Special Issue# 2.

Bosetti, V., Carraro, C., De Cian, E., Massetti, E., & Tavoni, M. (2012). *Incentives and stability of international climate coalitions: An integrated assessment*. FEEM Working Paper No. 97.2011

Bosetti, V., Heugues, M., & Tavoni, A. (2017). Luring others into climate action: coalition formation games with threshold and spillover effects. *Oxford Economic Papers*, *69*(2), 410-431.

Brannlund, R., & Persson, L. (2012). To tax, or not to tax: preferences for climate policy attributes. *Climate Policy*, *12*(6), 704-721.

Bristow, A. L., Wardman, M., Zanni, A. M., & Chintakayala, P. K. (2010). Public acceptability of personal carbon trading and carbon tax. *Ecological Economics*, *69*(9), 1824-1837.

Burke, M. A., & Young, H. P. (2011). Social norms. In K. J. Arrow, & M. D. Intriligator, *Handbook of Social Economics* (Vol. 1, pp. 311-338). North-Holland.

Cabrales, A., & Ponti, G. (2015). Social preferences. In P. Branas-Garza, A. Cabrales, *Experimental Economics* (Vol. 1, pp. 87-104). Palgrave Macmillan, London.

Cabrales, A., & Gottardi, P. (2022). *Network formation and heterogeneous risks* [in preparation].

Cabrales, A., García, M., Muñoz, D. R., & Sánchez, A. (2022). *The Interactions of Social Norms about Climate Change: Science, Institutions and Economics* [article submitted for publication].

Cai, B., Cameron, T. A., & Gerdes, G. R. (2010). Distributional preferences and the incidence of costs and benefits in climate change policy. *Environmental and Resource Economics*, *46*(4), 429-458.

Calzolari, G., Casari, M., & Ghidoni, R. (2018). Carbon is forever: A climate change experiment on cooperation. *Journal of Environmental Economics and Management*, *92*, 169-184.

Carattini, S., Baranzini, A., Thalmann, P., Varone, F., & Vöhringer, F. (2017). Green taxes in a post-Paris world: are millions of nays inevitable?. *Environmental and Resource Economics*, *68*(1), 97-128.

Carattini, S., Carvalho, M., & Fankhauser, S. (2018). Overcoming public resistance to carbon taxes. *Wiley Interdisciplinary Reviews: Climate Change*, *9*(5), e531.

Carattini, S., Levin, S., & Tavoni, A. (2020). Cooperation in the climate commons. *Review of Environmental Economics and Policy*, *13*(2).

Carlsson, F., & Johansson-Stenman, O. (2012). Behavioral economics and environmental policy. *Annual Review of Resource Economics*, *4*(1), 75-99.

Carraro, C., & Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, *52*(3), 309-328.

Chander, P., & Tulkens, H. (2006). A core-theoretic solution for the design of cooperative agreements on transfrontier pollution. In P. Chander, J. Drèze, C. Knox Lovell, J. Mintz, *Public goods, environmental externalities and fiscal competition* (pp. 176-193). Springer.

Chander, P. (2007). The gamma-core and coalition formation. *International Journal of Game Theory*, *35*(4), 539-556.

Chetty, R., Looney, A. & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review*, *99*(4), 1145-1177.

Cialani, C., & Mortazavi, R. (2018). Household and industrial electricity demand in Europe. *Energy Policy*, *122*, 592-600.

Cole, H. L., Mailath, G. J., & Postlewaite, A. (1992). Social norms, savings behavior, and growth. *Journal of Political Economy*, *100*(6), 1092-1125.

Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., & Tavoni, A. (2015). On the provision of public goods with probabilistic and ambiguous thresholds. *Environmental and Resource Economics*, *61*(3), 365-383.

Dannenberg, A., & Barrett, S. (2018). Cooperating to avoid catastrophe. *Nature Human Behaviour*, *2*(7), 435-437.

Doherty, K. L., & Webler, T. N. (2016). Social norms and efficacy beliefs drive the alarmed segment's public-sphere climate actions. *Nature Climate Change*, *6*(9), 879-884.

Dong, Y., Ma, S., Zhang, B., Wang, W. X., & Pacheco, J. M. (2021). Financial incentives to poor countries promote net emissions reductions in multilateral climate agreements. *One Earth*, *4*(8), 1141-1149.

Dreber, A., & Nowak, M. A. (2008). Gambling for global goods. *Proceedings of the National Academy of Sciences*, *105*(7), 2261-2262.

Drews, S. & Van den Bergh, J. C. (2016). What explains public support for climate policies? A review of empirical and experimental studies. *Climate Policy*, *16*(7), 855–876.

Eichhorn, J., Molthof, L. & Nicke, S. (2020). *From climate change awareness to climate crisis action: Public perceptions in Europe and the United States*. Open Society Foundations. https://www.opensocietyfoundations.org/publications/from-climate-change-awareness-to-climate-crisis-action (last accessed September 2, 2021).

Elster, J. (1989). Social norms and economic theory. *Journal of Economic Perspectives*, *3*(4), 99-117.

Faunce, T.A. (2012). Global public goods. In R. Chadwick, *Encyclopedia of Applied Ethics, Second Edition*. Elsevier.

Fehr, E., & Gächter, S. (2000). Cooperation and punishment in public goods experiments. *American Economic Review*, *90*(4), 980-994.

Fehr, E., & Leibbrandt, A. (2011). A field study on cooperativeness and impatience in the tragedy of the commons. *Journal of Public Economics*, *95*(9-10), 1144-1155.

Fehr, E., & Schurtenberger, I. (2018). Normative foundations of human cooperation. *Nature Human Behaviour*, *2*(7), 458–468.

Fischbacher, U., & Gächter, S. (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review*, *100*(1), 541-56.

Finus, M., Altamirano-Cabrera, J. C., & Van Ierland, E. C. (2005). The effect of membership rules and voting schemes on the success of international climate agreements. *Public Choice*, *125*(1), 95-127.

Gevrek, Z. E., & Uyduranoglu, A. (2015). Public preferences for carbon tax attributes. *Ecological Economics*, *118*, 186–197.

Ghesla, C., Grieder, M., & Schubert, R. (2020). Nudging the poor and the rich–A field study on the distributional effects of green electricity defaults. *Energy Economics*, *86*, 104616.

Ghidoni, R., Calzolari, G., & Casari, M. (2017). Climate change: Behavioral responses from extreme events and delayed damages. *Energy Economics*, *68*, 103-115.

Golman, R., Hagmann, D., & Loewenstein, G. (2017). Information avoidance. *Journal of Economic Literature*, *55*(1), 96-135.

Graziano, M., & Gillingham, K. (2015). Spatial patterns of solar photovoltaic system adoption: the influence of neighbors and the built environment. *Journal of Economic Geography*, *15*(4), 815-839.

Hagel, K., Abou Chakra, M., Bauer, B., & Traulsen, A. (2016). Which risk scenarios can drive the emergence of costly cooperation?. *Scientific Reports*, *6*(1), 1-9.

Hall, P. A., & Fong, G. T. (2007). Temporal self-regulation theory: A model for individual health behavior. *Health Psychology Review*, *1*(1), 6-52.

Halpern, D. (2015). *Inside the nudge unit: How small changes can make a big difference*. Random House.

Hardin, G. (1968). The Tragedy of the Commons: The population problem has no technical solution; it requires a fundamental extension in morality. *Science*, *162*(3859), 1243-1248.

Heath, Y., & Gifford, R. (2002). Extending the theory of planned behavior: Predicting the use of public transportation. *Journal of Applied Social Psychology*, *32*(10), 2154-2189.

Heitzig, J., & Kornek, U. (2018). Bottom-up linking of carbon markets under far-sighted cap coordination and reversibility. *Nature Climate Change*, *8*(3), 204-209.

Heres, D. R., Kallbekken, S., & Galarraga, I. (2017). The role of budgetary information in the preference for externality-correcting subsidies over taxes: a lab experiment on public support. *Environmental and Resource Economics*, *66*(1), 1-15.

Herrmann, B., Thoni, C., & Gachter, S. (2008). Antisocial Punishment Across Societies. *Science*, *319*(5868), 1362-1367.

Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and boosting: Steering or empowering good decisions. *Perspectives on Psychological Science*, *12*(6), 973-986.

Hilbe, C., Abou Chakra, M., Altrock, P. M., & Traulsen, A. (2013). The evolution of strategic timing in collective-risk dilemmas. *PloS ONE*, *8*(6), e66490.

Hortaçsu, A., Madanizadeh, S. A., & Puller, S. L. (2017). Power to choose? An analysis of consumer inertia in the residential electricity market. *American Economic Journal: Economic Policy*, *9*(4), 192-226.

Hsu, S. L., Walters, J., & Purgas, A. (2008). Pollution tax heuristics: An empirical study of willingness to pay higher gasoline taxes. *Energy Policy*, *36*(9), 3612–3619.

Huber, R. A., Anderson, B., & Bernauer, T. (2018). Can social norm interventions promote voluntary pro environmental action?. *Environmental Science and Policy*, *89*, 231-246.

IPCC, Intergovernmental Panel on Climate Change. (2018). *Global Warming of 1.5° C: An IPCC Special Report on the Impacts of Global Warming of 1.5° C Above Pre-industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty.*

IPCC, Intergovernmental Panel on Climate Change. (2019). *Climate Change and Land: An IPCC Special Report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.*

IPCC, Intergovernmental Panel on Climate Change. (2022). Technical Summary. In H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, and B. Rama, *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. In Press.

Jacquet, J., Hagel, K., Hauert, C., Marotzke, J., Röhl, T., & Milinski, M. (2013). Intra-and intergenerational discounting in the climate game. *Nature Climate Change*, *3*(12), 1025-1028.

Kaiser, F. G., & Gutscher, H. (2003). The proposition of a general version of the theory of planned behavior: Predicting ecological behavior. *Journal of Applied Social Psychology*, *33*(3), 586-603.

Kallbekken, S., & Aasen, M. (2010). The demand for earmarking: Results from a focus group study. *Ecological Economics*, *69*(11), 2183–2190.

Kallbekken, S., & Sælen, H. (2011). Public acceptance for environmental taxes: Self-interest, environmental and distributional concerns. *Energy Policy*, *39*(5), 2966–2973.

Kalmi, P., Trotta, G., & Kažukauskas, A. (2021). Energy–related financial literacy and electricity consumption: Survey–based evidence from Finland. *Journal of Consumer Affairs*, *55*(3), 1062-1089.

Karatayev, V. A., Vasconcelos, V. V., Lafuite, A. S., Levin, S. A., Bauch, C. T., & Anand, M. (2021). A well-timed shift from local to global agreements accelerates climate change mitigation. *Nature Communications*, *12*(1), 1-7.

Kaufmann, C., Weber, M., & Haisley, E. (2013). The role of experience sampling and graphical displays on one's investment risk appetite. *Management Science*, *59*(2), 323-340.

Kimbrough, E. O., & Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association*, *14*(3), 608-638.

Klok, J., Larsen, A., Dahl, A., & Hansen, K. (2006). Ecological tax reform in Denmark: History and social acceptability. *Energy Policy*, *34*(8), 905–916.

Kölle, F. (2015). Heterogeneity and cooperation: The role of capability and valuation on public goods provision. *Journal of Economic Behavior and Organization*, *109*, 120-134.

Kölle, F., & Lauer, T. (2020). *Understanding cooperation in an intertemporal context* (No. 046). ECONtribute Discussion Paper.

Lázaro Touza, L., González Enríquez, C. & Escribano, G. (2019). *Los españoles ante el cambio climático*. Real Instituto Elcano. http://www.realinstitutoelcano.org/wps/portal/rielcano_es/contenido?WCM_GLOBAL_CONTEXT=/elcano/elcano_es/zonas_es/cambio-climatico/encuesta-espanoles-ante-cambio-climatico-sep-2019 (last accessed September 2, 2021).

Lusardi, A., Samek, A., Kapteyn, A., Glinert, L., Hung, A., & Heinberg, A. (2017). Visual tools and narratives: New ways to improve financial literacy. *Journal of Pension Economics and Finance*, *16*(3), 297-323.

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, *14*(3), 115-136.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S.L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M.I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J.B.R., Maycock, T.K., Waterfield, T., Yelekçi, O., Yu, R., & Zhou, B. (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://www.ipcc.ch/report/ar6/wg1/#TS (last accessed September 2, 2021)

McEvoy, D. M., & Stranlund, J. K. (2016). *Inequality Aversion and Coalition Formation*. Working Papers 16-09, Department of Economics, Appalachian State University.

Mehling, M. A., Van Asselt, H., Das, K., Droege, S., & Verkuijl, C. (2019). Designing border carbon adjustments for enhanced climate action. *American Journal of International Law*, *113*(3), 433-481.

Milinski, M., Sommerfeld, R. D., Krambeck, H. J., Reed, F. A., & Marotzke, J. (2008). The collective-risk social dilemma and the prevention of simulated dangerous climate change. *Proceedings of the National Academy of Sciences*, *105*(7), 2291-2294.

Muldoon, R., Lisciandra, C., Bicchieri, C., Hartmann, S., & Sprenger, J. (2014). On the emergence of descriptive norms. *Politics, Philosophy and Economics*, *13*(1), 3-22.

Nilsson, A., von Borgstede, C., & Biel, A. (2004). Willingness to accept climate change strategies: The effect of values and norms. *Journal of Environmental Psychology*, *24*(3), 267-277.

Nordhaus, W. (2015). Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review*, *105*(4), 1339-1370.

Nordhaus, W. (2021). Dynamic climate clubs: On the effectiveness of incentives in global climate agreements. *Proceedings of the National Academy of Sciences*, *118*(45), e2109988118.

Nordlund, A. M., & Garvill, J. (2003). Effects of values, problem awareness, and personal norm on willingness to reduce personal car use. *Journal of Environmental Psychology*, *23*(4), 339-347.

Nyborg, K., Anderies, J. M., Dannenberg, A., Lindahl, T., Schill, C., Schlüter, M., Adger, W.N., Arrow, K.J., Barrett, S., Carpenter, S., Stuart Chapin III, F., Crépin, A-S., Gretchen D., Ehrlich P., Folke, C., Jager, W., Kautsky, N., Levin, S.A., Madsen, O.J. ... De Zeeuw, A. (2016). Social norms as solutions. *Science*, *354*(6308), 42-43.

Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard University Press.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives*, *14*(3), 137-158.

Pacheco, J. M., Santos, F. C., Souza, M. O., & Skyrms, B. (2009). Evolutionary dynamics of collective action in N-person stag hunt dilemmas. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1655), 315-321.

Paluck, E. L. (2009a). Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda. *Journal of Personality and Social Psychology*, *96*(3), 574.

Paluck, E. L. (2009b). What's in a norm? Sources and processes of norm change. *Journal of Personality and Social Psychology*, *96*(3), 594–600.

Pollitt, M. G., & Shaorshadze, I. (2013). The role of behavioural economics in energy and climate policy. In R. Fouquet, *Handbook on Energy and Climate Change* (pp. 523-546). Edward Elgar Publishing.

Ramos Muñoz, D., Cabrales, A., & Sanchez, A. (2022). *Central banks and climate change. Fit, opportunity and suitability in the law and beyond*. European Banking Institute Working Paper Series 2022 - No. 119.

Rebonato, R. (2012). *Taking liberties: A critical examination of libertarian paternalism*. Palgrave Macmillan.

Rode, J., & Weber, A. (2016). Does localized imitation drive technology adoption? A case study on rooftop photovoltaic systems in Germany. *Journal of Environmental Economics and Management*, *78*, 38-48.

Rogna, M. (2019). Coalition formation and bargaining protocols: a review of the literature. *Journal of Economic Surveys*, *33*(1), 226-251.

Sandler, T., & Tschirhart, J. T. (1980). The economic theory of clubs: An evaluative survey. *Journal of Economic Literature*, *18*(4), 1481-1521.

Schmidt, K. M., & Ockenfels, A. (2021). Focusing climate negotiations on a uniform common commitment can promote cooperation. *Proceedings of the National Academy of Sciences*, *118*(11), e2013070118.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429-434.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General*, *130*(3), 380-400.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*(1), 99-118.

Sparkman, G., Howe, L., & Walton, G. (2021). How social norms are often a barrier to addressing climate change but can be part of the solution. *Behavioural Public Policy*, *5*(4), 528-555.

Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science*, *333*(6048), 1393-1400.

Sturm, B., & Weimann, J. (2006). Experiments in environmental economics and some close relatives. *Journal of Economic Surveys*, *20*(3), 419-457.

Svenningsen, L. S., & Thorsen, B. J. (2020). Preferences for distributional impacts of climate policy. *Environmental and Resource Economics*, *75*(1), 1-24.

Szekely, A., Lipari, F., Antonioni, A., Paolucci, M., Sánchez, A., Tummolini, L., & Andrighetto, G. (2021). Evidence from a long-term experiment that collective risks change social norms and promote cooperation. *Nature Communications*, *12*(1), 1-7.

Tagliapietra, S., & Wolff, G. B. (2021). Form a climate club: United States, European Union and China. Comment. *Nature*, *591*(7851), 526-528.

Tavoni, A., Dannenberg, A., Kallis, G., & Löschel, A. (2011). Inequality, communication, and the avoidance of disastrous climate change in a public goods game. *Proceedings of the National Academy of Sciences*, *108*(29), 11825-11829.

Thaler, R., & Sunstein, C.R. (2008). *Nudge: Improving decisions about health, wealth and happiness*. Penguin.

Todorov, A., & Mandisodza, A. N. (2004). Public opinion on foreign policy: The multilateral public that perceives itself as unilateral. *Public Opinion Quarterly*, *68*(3), 323-348.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124-1131.

Van der Linden, S. L., Leiserowitz, A. A., Feinberg, G. D., & Maibach, E. W. (2015). The scientific consensus on climate change as a gateway belief: Experimental evidence. *PloS ONE*, *10*(2), e0118489.

Van der Werff, E., Steg, L., & Keizer, K. (2014). I am what I am, by looking past the present: The influence of biospheric values and past behavior on environmental self-identity. *Environment and Behavior*, *46*(5), 626-657.

Vogt, C. (2016). Climate coalition formation when players are heterogeneous and inequality averse. *Environmental and Resource Economics*, *65*(1), 33-59.

Weber, E. U. (2006). Experience-based and description-based perceptions of long-term risk: Why global warming does not scare us (yet). *Climatic Change*, *77*(1), 103-120.

Weikard, H.-P., Finus, M., & Altamirano-Cabrera, J. C. (2006). The impact of surplus sharing on the stability of international climate agreements. *Oxford Economic Papers, 58*(2), 209-232.

Weitzman, M. L. (2014). Can negotiating a uniform carbon price help to internalize the global warming externality?. *Journal of the Association of Environmental and Resource Economists*, *1*(1/2), 29-49.

Winchester, N., Paltsev, S., & Reilly, J. M. (2011). Will border carbon adjustments work?. *The BE Journal of Economic Analysis & Policy*, *11*(1).

Yang, Z. (1999). Should the North make unilateral technology transfers to the South?: North–South cooperation and conflicts in responses to global climate change. *Resource and Energy Economics*, *21*(1), 67-87.

Yang, Z. (2008). *Strategic bargaining and cooperation in greenhouse gas mitigations: an integrated assessment modeling approach*. MIT Press.

Young, H. P. (2015). The evolution of social norms. *Annual Review of Economics*, *7*(1), 359-387.

Zhu, X., Li, L., Zhou, K., Zhang, X., & Yang, S. (2018). A meta-analysis on the price elasticity and income elasticity of residential electricity demand. *Journal of Cleaner Production*, *201*, 169-177.