



A new two-nested-game approach: linking micro- and macro-scales in international environmental agreements

Tiziano Distefano¹ · Simone D'Alessandro¹

Accepted: 2 February 2021 / Published online: 19 February 2021
© The Author(s) 2021

Abstract

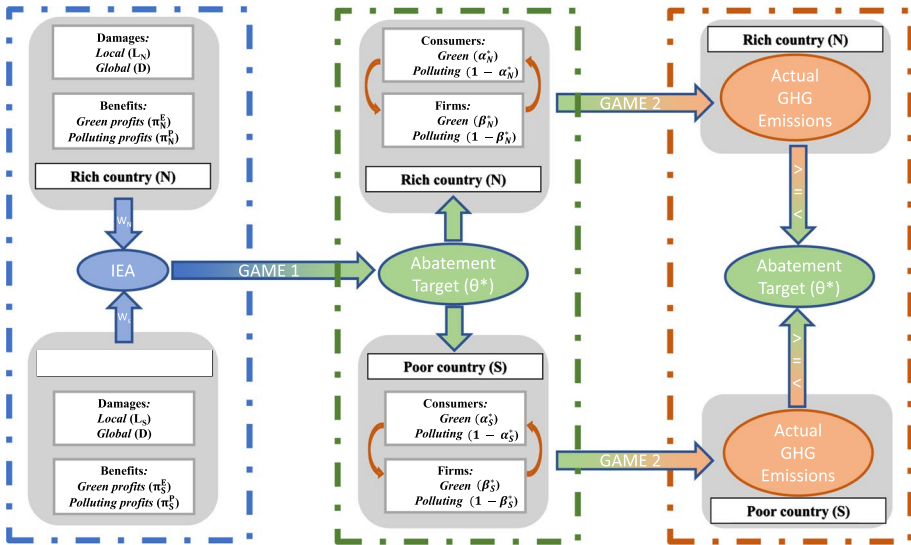
This study investigates the effectiveness of international environmental agreements (*IEAs*) and how it might be affected by the development of pro-environmental behaviour among households and firms. We propose a new framework based on a *two-nested-game* approach composed by: (1) a one-shot game with two *asymmetric* countries that negotiate the international abatement target, and (2) an evolutionary game which describes the economic structure resulting from agents' interactions. These two games are nested because the initial economic structure determines the welfare of each country, and thus the outcome of Game 1 which, in turn, is embedded in Game 2, modifies the agents' pay-off and the economic structure thereof. Numerical simulation outcomes suggest three key messages. First, we find that global solutions do not automatically produce the expected effects irrespective of any free-riding assumption. Second, extreme climate risks might not lead to a high abatement target in the event of marked cross-country inequality. Third, adverse consumers' environmental attitudes might hamper the success of an *IEA*. The above observations entail that governments should not simply impose environmental laws. Rather, top-down policies and bottom-up interventions should be coordinated; otherwise, they might fail if undertaken in isolation.

✉ Tiziano Distefano
tiziano.distefano@unipi.it

Simone D'Alessandro
simone.dalessandro@unipi.it

¹ Department of Economics and Management, University of Pisa, Via C. Ridolfi, 10, 56124 Pisa, Italy

Graphic abstract



Keywords Inequality · Evolutionary games · Environment and growth · Climate change · Environmental attitude

JEL Classifications B52 · D10 · D63 · O44 · Q54

1 Introduction

Anthropogenic climate change and trans-boundary pollution have long justified the emergence of international environmental agreements (henceforth *IEAs*). There are increasing efforts worldwide to find joint solutions to curb greenhouse gas (GHG) emissions (and to save the ecosystems in general) while ensuring socio-economic security, as shown by the definition of the seventeen Sustainable Development Goals.¹ The issue of stability and effectiveness of *IEAs* has attracted growing attention from the academic literature in recent decades.

Since the seminal paper of Barrett (1994), the economic literature has approached this issue by modelling the optimal *IEA* coalition size within the framework of game theory (see Marrouch et al. 2016, for a review). The general consensus is that a global agreement on emissions reduction is not feasible (Yang 2017) and that only small coalitions are stable (Gelves and McGinty 2016). On the contrary, the actual number of parties to climatic *IEAs* has increased over time, from 84 countries that signed the first ratification of the Kyoto

¹ See <https://sdgs.un.org/goals>.

Protocol (COP3) in 1999 to 200 countries participating in the last COP25 held in Madrid in 2019.²

Beyond that, few key players actually determine the structure and success of self-enforcing *IEAs* (Finus et al. 2009). A case in point was the US reversal from a leadership role in forming climate coalitions in the lead-up to the Paris Agreement (COP 23 in 2015), to its announcement of a withdrawal from the same treaty in 2017 (Zhang et al. 2017). This situation reflects the so-called breadth and depth dilemmas (Keohane and Victor 2011; Hannam et al. 2017), i.e. the trade-off between breadth (wanting many actors) and depth (wanting strength of action/mitigation) agreements.³ Despite many theoretical and empirical advances, one under-explored aspect is the role of domestic economic structures on the outcome of climate agreements (and their effectiveness thereof). The current contribution aims to fill this gap.

Given this premise, we aim to shed a light on the following questions: 1) To what extent does the *country-specific economic structure* determines the success or failure of *IEAs*? 2) How do global *inequalities* and economic growth affect the road towards a low-carbon transition? 3) Are *local* initiatives able to boost pro-environmental behaviours or do they hamper the success of environmental laws? To answer these questions, we propose a novel framework, based on a *two-nested-game* approach that covers alternative economic structures, cross-country inequalities, and different consumer environmental attitudes. We propose a model that analyses the linkages between the macro-scale (international agreements) and the micro-scale (firm and consumer choices) to evaluate the actual effectiveness of *IEAs*. Thus, environmental performances that differ from what is ratified—i.e. failure, fulfilment or over-performance (i.e. when CO₂ reductions exceed than what is agreed)—are explained by *endogenous* dynamics determined by both the stringency of the targets defined during the *IEAs* and the interactions among economic agents within each country.

This approach challenges some of the main (often hidden) assumptions behind the available studies, namely:

- a1* the ratified agreement will *automatically* yield, in the real economy, the planned effect;
- a2* the failure to comply with the *IEA* is only explained in terms of *free-riding* behaviour;
- a3* *local* initiatives are ineffective: people submit to environmental legislation without making any voluntary effort to support or hinder the introduction of such environmental laws.

Assumption *a1* appears in contrast with the fact that, despite the ratification of several climatic *IEAs*, GHG emissions are still increasing at the global level (Testa et al. 2016). Moreover, there is considerable cross-country heterogeneity in environmental performance, in terms of the difference between what is signed in *IEAs* and the actual level of emissions.⁴

² It followed up the Paris Agreement, designed to replace the Kyoto Protocol, that was signed by 197 countries and ratified by 187 as of November 2019.

³ For instance, Paris COP21 is framed as an umbrella of small and large agreements, which allow for small and big coalitions to benefit from the framework and conventions established by the treaty.

⁴ For instance, some Kyoto participants were well above their targets while others fell well below. There are some successful examples of emission reductions, such as France, Italy, Germany and the UK (see Olivier et al. 2013). We do not focus on the carbon leakage issue.

The assumption $a2$ imposes a free-riding behaviour—i.e. despite the global benefits of reducing GHG emissions, no agent has the incentive to reduce his/her own burden—neglecting the possibility of voluntary initiatives. On the contrary, even developing countries decided to implement national environmental policies although they were exonerated from emission controls. India approved a fund of US\$ 1,371 million to boost energy efficiency and develop clean technologies and electrification, while since 2014 China has deployed more solar and wind capacity than any country worldwide, adding around 17.5 Gigawatts of solar energy by the end of 2015 alone.⁵

The last assumption ($a3$) excludes bottom-up initiatives. In reality, many environmental NGOs (e.g. Greenpeace and WWF), climate campaigners (“Fridays for future”, the school strike movement) and local communities have put pressure on governments (e.g. the EU’s Covenant of Mayors initiative) in order to actively protect and preserve ecological systems (Nasiritousi et al. 2016). On the other hand, increasing economic disparities and job instability within national boundaries (Alvaredo et al. 2018) have led to social unrest against the introduction of green laws (Ponticelli and Voth 2020) such as the protests of the “Gilets Jaunes” after a carbon tax was introduced by the French government in November 2018.⁶

This study is structured as follows. Section 2 briefly explains the methodology and the logic of the two games. Section 3 describes the one-shot 2x2 asymmetric IEA game, and Sect. 4 deals with the evolutionary interactions between households and firms. Section 5 presents and discusses the results of alternative scenarios based on numerical simulations, defined by a handy Maple algorithm. Finally, Sect. 6 draws the main conclusions and policy implications.

2 Methods

Before moving onto the mathematical details, it is worth explaining the logic behind our model in order to clarify the interpretations thereof. Figure 1 shows a schematic representation of the step-by-step procedure proposed. We develop a model which integrates the results from two nested games.⁷ Game 1 analyses the results (at the macro-scale) of a one-shot 2x2 game, where two *asymmetric* countries⁸—representative of the richer (N) and poorer (S) regions, respectively (e.g. Anand 2017)—bargain over the share of clean energy (θ) to attain given emission abatement targets.⁹ Each country differs in terms of income level and technological development which determines the attainable value of the welfare function (W).

Game 2 embeds the outcome of Game 1 (i.e. θ) that modifies the pay-off of firms and consumers. Note that Game 2 models the evolutionary dynamic (at the micro-scale) of

⁵ See <https://ggi.org/project/india-green-growth-through-ndc/> and <https://ggi.org/project/china-global-and-regional-knowledge-sharing/>.

⁶ See <https://jacobinmag.com/2018/11/yellow-vests-fuel-prices-france-protests>.

⁷ Note that, for the sake of clarity, we focus on the impact of international laws (i.e. top-down) on domestic economies, while the inclusion of the influence of bottom-up initiatives on international agreements lies outside the scope of the current study.

⁸ A real case example is the meeting between the two biggest polluters, the USA and China, held in November 2014.

⁹ We indirectly capture the agreement on abatement targets via renewable resources; thus, for instance, if a country wants to halve its carbon emissions, we assume, as explained in Sect. 2, that it must have a share of green firms equal to 50% (i.e. $\theta = 0.5$).

the country's economic structure represented by the interaction of the economic decisions of consumers and firms, who decide whether or not to be environmentally friendly. Their interactions determine *endogenously* the actual level of carbon emissions, which can differ from what is agreed in the bilateral *IEA*.

The *timing* of the whole model is as follows: each country recognizes the importance of transboundary effects of carbon emissions and then decides to bargain (i.e. Game 1 on bilateral *IEA*) to ratify an international environmental standard (θ). This target is conceived as the minimum share of clean production, with respect to industrial profits (or GDP), thus reflecting a decision on the energy mix of the production process. Indeed, the amount of renewable resources increases with θ because firms have to install new green technologies to curb polluting emissions.

The treaty is enforced by national law and so θ is embodied in the agents' pay-off. *Game 2* determines the evolutionary stable state(s) depending on the evolutionary strategic interactions between consumers and firms. The emissions in equilibrium are then compared with the level of emissions ratified in the *IEA* (right-hand panel of Fig. 1). Note that we do not consider any feedback from *Game 2* to *Game 1*. Although this might represent a limitation, the main purpose of our study is to assess to what extent the economic structure influences the actual level of emissions, considering that the government has no perfect information about the complexity of the underlying economic system (Hatase and Matsubayashi 2019). Hence, our approach explains under what (micro- and macro-) conditions environmental performance might differ from what was ratified during the *IEA*.

3 Bilateral *IEA* (game 1)

Each country $i = \{N, S\}$ has a welfare objective function defined as:¹⁰

$$W_i(\theta_i, \pi_i) = B_i(\theta_i, \pi_i) - a_i \cdot D(\theta, \pi), \quad (1)$$

where the parameter $0 \leq a_i \leq 1$ represents the country-specific *climate risk*, i.e. the probability of undergoing economic losses due to climate change induced by global pollution (Cai et al. 2013). As is commonly found in the literature on static simultaneous move games (see Marrouch et al. 2016, p.249), the welfare function is defined as the difference between the net economic benefits of production minus the estimated (monetary) costs of environmental damage. We define the net benefits (B_i) as the difference between the overall country-specific economic production (Π_i) and the *local* damage (L_i , i.e. within national boundaries), expressed as a quadratic function (see Pavlova and de Zeeuw 2013). Hence, even in the absence of any *IEA* a country might find it worth introducing environmental laws to tackle environmental deterioration and health problems due to industrial discharges of national polluting firms. Total country-specific industrial profits (Π_i) are determined by the profits of clean (π_i^E) and polluting (π_i^P) profits weighted by the share of green (θ) and brown ($1 - \theta$) firms, respectively. Note that a one-to-one relationship is assumed between production and emissions such that π_p denotes both production and emissions (see Pavlova and de Zeeuw 2013). Namely:

¹⁰ Note that, in line with the current literature, W_i does not include consumers' utility for two reasons: (i) we want to compare our results within a classical framework to understand the contribution of Game 1 and (ii) consumers' utility does not affect the welfare in equilibrium.

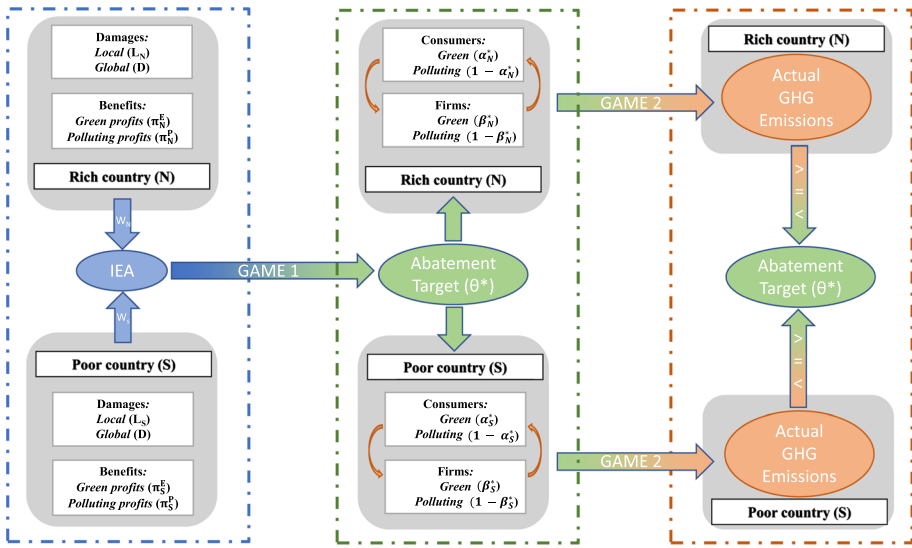


Fig. 1 Macro-view of the model structure. Schematic representation of the logic and the sequential phases composing the *two nested* games. On the left (blue panel) the structure of each country—i.e. *benefits* from green (π^E) and brown (π^P) profits and *damages* due to local (L) and global (D) pollution—determines the Welfare function of the rich (W_N) and poor (W_S) country. Then, Game 1 (green panel) determines the outcome of the IEA, i.e. the optimal abatement target (θ^*) that will be implemented in each country by law. The environmental law will affect the firms and consumers pay-off. In Game 2 they interact until an evolutionary equilibrium (α^*, β^*) emerges. Finally (orange panel), the country’s actual GHG emissions will be compared with what was ratified in the IEA, being greater, equal or lower

$$B_i(\theta_i, \pi_i) = b_i[d_i \cdot \Pi_i(\theta_i, \pi_i) - L_i(\theta_i, \pi_i)], \tag{2}$$

$$\Pi_i(\theta_i, \pi_i) = \theta_i \cdot \pi_i^E + (1 - \theta_i) \cdot \pi_i^P, \tag{3}$$

$$L_i(\theta_i, \pi_i) = \frac{[(1 - \theta_i) \cdot \pi_i^P]^2}{2}, \tag{4}$$

where the parameter $0 \leq b_i \leq 1$ is the overall opportunity cost, while $0 \leq d_i \leq 1$ is the marginal industrial benefit. The damage to country i from global pollution (D_i) is a quadratic function of the global emissions¹¹—given by the sum of domestic and foreign pollution—representing a proxy of the potential damage caused by extreme climate events (Coronese et al. 2019). Namely

¹¹ Note that the quadratic climate damage function is widely applied in the literature, such as in the DICE, ICAM and MERGE models. Detailed investigation of the actual form of the climate damage function goes beyond the scope of the present paper and would be too complex given the high level of uncertainty and difficulties in the monetary valuation involved (see Richard 1995, for a review). For this reason, we retain the most common definition because the main interest is to evaluate the influence of the domestic economic structure on the success of IEAs.

$$D(\theta, \pi) = \left(\sum_{k=\{i,j\}} (1 - \theta_k) \cdot \pi_k^P \right)^2 \tag{5}$$

Note that using a linear-quadratic framework allows us to avoid discrete solutions (pollute or not) and then to provide a more realistic representation of the abatement target ratified during the *IEA*.

3.1 Agreement formation in the international setting

The decision rule to find the optimal single agreement is based on the “smallest common denominator” (*SCD*) to ensure the external stability of the *IEA*. The *SCD* rule implies that a compromise is found on the lowest environmental target: if $\theta_N > \theta_S$, then the Nash equilibrium is $\theta^* = \theta_S$, and the reverse (Finus 2002). Although one would expect proposals to be strategically motivated, it is possible to demonstrate that the *SCD* decision rule is immune to strategic offers; rather, it is a best reply and a Nash equilibrium (Endres and Finus 1999, pp. 539–540).

From Eqs. (1) and (2), it derives that each country has an incentive to implement environmental laws even in the absence of an *IEA* ($a_i = 0$) to tackle local environmental damage from polluting production, as formally discussed in “Appendix” A (see the business-as-usual case). For the sake of clarity, here we only expose the analytical results from the *IEA*, viz. when global externalities are recognized ($a_i > 0$). The *non-cooperative* Nash equilibrium is given by the maximization of Eq. (1) with respect to θ_i . Namely

$$\begin{aligned} \theta_i = & K_i \{ \pi_j^P (b_j b_i + a_j b_i) [(\pi_i^P)^2 - d_i (\pi_i^P - \pi_i^E)] \\ & + \pi_i^P \cdot a_i b_j \cdot [\pi_i^P \pi_j^P + d_j (\pi_j^P - \pi_j^E)] \}, \end{aligned} \tag{6}$$

where

$$K_i = [\pi_j^P (\pi_i^P)^2 \cdot (b_j b_i + a_j b_i + a_i b_j)]^{-1}, \tag{7}$$

with $(j, i) = \{N, S\}$ and $j \neq i$.

Given the high nonlinearity of these solutions, it is not possible to establish a simple relationship between the optimal level of environmental standards and the parameters and variables involved. For this reason, Sect. 5 performs numerical simulations, while in “Appendix” A.2 there are the *ceteris paribus* analytical discussions (propositions and proofs) about the conditions that determine the bottleneck of the *IEA*.

4 Evolutionary (micro-)economic structure (game 2)

In the second step, we construct a dynamic model of the process by which the proportions of various strategies in a population evolve (Weibull 1997). To analyse the evolutionary dynamics governing transitions between the two conventions (Green–Green, Carbon Economy), let us assume a two-person two-strategy ($s = \{E, P\}$) game in a large population of individuals subdivided into two groups—namely households (\mathcal{H}) and firms (\mathcal{F})—the members of which are randomly matched to interact in a non-cooperative game with the members of the other group. Individuals’ best-response play is based on a single-period

memory: they maximize their expected pay-off based on the distribution of the population in the previous period (Bowles 2009; Belloc and Bowles 2013). Both populations are normalized to unit size, so we refer equivalently to the numbers of players and to the fraction of the population.¹² The application of evolutionary games to study the link between economics and the environment is gaining attention because of its ability to disentangle complex dynamics (e.g. Antoci et al. 2019).

We assume that when people with different strategies are matched they do not sign any contract; in other words, green consumers do not want to buy polluting goods and *vice versa*.¹³ The dynamic evolution of the proportion of ecological households and firms (α and β , respectively) follows the so-called replicator dynamic (Santos and Pacheco 2011) :

$$\dot{\alpha}_t = \alpha_t \cdot (1 - \alpha_t) \cdot [H_{E,t} - H_{P,t}] \tag{8}$$

$$\dot{\beta}_t = \beta_t \cdot (1 - \beta_t) \cdot [F_{E,t} - F_{P,t}] \tag{9}$$

where $H_{s,t}$ and $F_{s,t}$ are the expected fitness of households and firms' strategies, respectively, in period t . According to Eqs. (8) and (9), the percentage of green players increases if the pay-off given by the green strategy (E) is higher than what is expected when the polluting strategy (P) is played. The pay-off depends on the actions of the co-players and hence on the frequencies of the strategies within the population (Sigmund 2010). The more successful individuals will be "mimicked" by others, so that the share of individuals adopting a given strategy changes over time.

4.1 Households

The utility of a household h depends on its material pay-off and strategy (h_s). For the sake of simplicity, we assume that the consumption of each kind of commodity yields the same level of utility (u).¹⁴ The pay-off of the household is a piece-wise continuous function defined as:

$$h_{s,t} = \begin{cases} h_{E,t} = u - c(\theta - \beta_t), & \text{iff } \beta_t < \theta \text{ and } s = E \\ h_{P,t} = u - \delta(\theta - \beta_t), & \text{iff } \beta_t < \theta \text{ and } s = P \\ u, & \text{otherwise, with } s = \{E, P\}. \end{cases} \tag{10}$$

where $0 < \delta, c, u \leq 1$. Parameter $u > 0$ is the constant level of utility from consumption, independent of market conditions (not modelled here).¹⁵ Environment friendly households

¹² Note that this framework does not deal with interactions that take place among more than two individuals at the same time.

¹³ Following Weibull (1997, p. 34) we might consider that the zeros outside the diagonal are the results of pay-off normalization. Hence, even if we assume that agents get some constant positive pay-off when they play a different strategy, it would not affect the structure of Game 1. For simplicity, we neglect this possibility and we only assess the model with zeros outside the diagonal.

¹⁴ In other words, green and polluting goods are perfect substitutes because both goods are able, through their material characteristics, to satisfy consumer needs in the same manner. For instance, a consumer should be indifferent between an ecological home cleaning product and a chemical one, if they are both equally suitable for housecleaning.

¹⁵ Note that the utility does not depend on quantity because we assume that in each pair-wise matching the amount of exchange is constant.

bear an additional monetary cost (c) proportional to the difference between the environmental standard ($0 \leq \theta \leq 1$) and the share of green firms ($0 \leq \beta \leq 1$). This additional cost represents the willingness to finance the (start-up) of green production. In other words we include the observations that the more eco-friendly the product is, the higher the price that consumers must pay (Shi et al. 2019).¹⁶ On the other hand, polluting households incur, as a moral cost, the weight of negative public opinion regarding the possible damage arising from polluting consumption (δ). Thus, δ is a measure of the level of *environmental attitude* commonly found in society, which might also be interpreted as a proxy of the willingness to accept the imposition of green policies. In other words, if δ is low it can be interpreted as a form of resistance or social unrest against the environmental laws, while if high it might be indicative of higher environmental awareness. Notably, moral motivations have been proved to be an important driver of behaviour, especially when concerning environmental choices (Turaga et al. 2010; Steg 2016).

The share of green households, in equilibrium, depends on the gap between the environmental law θ and the actual proportion of green firms (β^*), while in the case of no environmental concerns—i.e. $\beta \geq \theta$ —the green household simply receives utility from consumption. The interested reader can find the complete analytical discussion, including proofs and propositions, in “Appendix” B.1.

4.2 Firms

Let us assume that the firm’s pay-off is a piece-wise continuous function defined as:

$$f_{s,t} = \begin{cases} f_{E,t} = \pi_E + \frac{c(\theta-\beta_t)\alpha}{\beta_t}, & \text{iff } \beta_t < \theta \text{ and } s = E, \\ f_{P,t} = \pi_P - \frac{\gamma(\theta-\beta_t)}{1-\beta_t}, & \text{iff } \beta_t < \theta \text{ and } s = P, \\ \Pi_s, & \text{otherwise, with } s = \{E, P\}. \end{cases} \quad (11)$$

where $0 < \pi_E < \pi_P < 1$ because we assume that the green technology is not already developed to be as efficient as the polluting one and thus the related profits are smaller (Goeschl and Perino 2017).¹⁷ This distinction is crucial to assess how the energy mix of production, at the country scale, determines the country’s environmental performance.

The coefficient $\gamma > 0$ is a multiplicative factor that measures the monetary cost of the difference $\theta - \beta_t$ which represents the gap between the actual and the normative share of clean production. The level of profits from production depends on the strategy (π_s); they can be interpreted as an average profit proportional to the market share of the belonging sector. A *green firm*¹⁸—other than profits—receives a premium-price which is equal to the total amount of extra costs (c), paid by each green household, multiplied by the share of green consumers (α). This amount decreases as the share of green firms approaches the environmental standards ($\beta \rightarrow \theta$). We assume that, in every period, the total amount paid by green consumers is equally shared among green firms. The premium price sustains and

¹⁶ Modelling the process of price formation and the mechanism of redistribution of this extra payment lies beyond the scope of the current study. As a matter of example, we report, among the several real-case initiatives, that many green start-ups get off the ground using crowd-funding sites, such as “FoodCycle”, which recycles food waste into nutritious meals for those in need.

¹⁷ The share of renewable sources in meeting global energy demand was about 10% in 2017, see <https://www.iea.org/topics/renewables/>.

¹⁸ See Battaglia et al. (2018) for a discussion of green jobs.

boosts the investments in new green start-ups to stimulate investments in green sectors (Bergset and Fichter 2015).

The *polluting* firm faces a cost ($\frac{\gamma(\theta-\beta_i)}{1-\beta_i}$) for local damage that depends on: the relation between the actual level of green production (β_i), the abatement target fixed by the government (θ), and the number of polluting firms ($1 - \beta_i$). Since polluting firms must jointly cover the cost of environmental damage, the total amount is determined by the percentage of polluting firms. We assume that the amount paid by polluting firms might be used by the government to restore the detrimental effects and to finance climate mitigation actions.

Given that the expected pay-off of firms is not linear, depending on households strategy, we develop in Sect. 5 numerical simulations to evaluate all the possible cases. The interested reader can find the complete analytical discussion, including proofs and propositions, in “Appendix” B.2.

5 Results

For the sake of simplicity, we focus on four key dimensions: climatic risk distribution (a), global economic development ($\Pi_N + \Pi_S$), cross-country technological inequality (Π_N/Π_S),¹⁹ and consumers’ environmental attitudes (δ). For a comparative analysis about the effect of different opportunity costs, see “Appendix” A.2.

Table 1 reports the setting of the values of all the parameters that characterize the different scenarios. Depending on global economic growth and cross-country inequality, we define four alternative *cases*. Note that in Game 1 we have 16 alternative sets of pay-offs since that for each of the four *IEA* cases (inequality/growth) there are other four alternative games depending on the cross-country climatic risk distribution (i.e. $a_L^N - a_L^S$ or $a_L^N - a_H^S$ or $a_H^N - a_L^S$ or $a_H^N - a_H^S$). To this, Game 2 adds other two alternative worlds where consumers might have either low or high environmental attitudes, making a total of 32 scenarios.

Numerical simulations enable reliable outcomes to be generated when many parameters vary at the same time, which would be impossible to assess analytically. This is crucial to clarify the relation between the target ratified in the *IEA* and the actual level of carbon pollution which depends on the country’s economic structure and its evolutionary path.²⁰ Our framework provides reliable scenarios able to capture the interaction between top-down policy decisions and the (micro-level) economic structure. Note that the proposed scenarios capture all the possible cases, thereby producing clear results despite the complex issues at hand. Indeed, any other modification of the parameters will generate outcomes somewhat in between those defined herein. This, together with the analytical discussion provided in the “Appendix”, also ensures robustness of the outcomes.

5.1 Optimal IEA target

Table 2 collects all of the solutions from Game 1, also specifying whether the poorer (S) or richer (N) country is the bottleneck of the agreement, viz. who proposes the lowest target

¹⁹ This distinction follows the literature on North-South inequality (Roberts and Parks 2006; Grasso 2011). Thus, unless otherwise stated, we use the term *inequality* to indicate the uneven level of economic development in terms of technological progress and GDP level.

²⁰ Numerical simulations were performed using a Maple algorithm that is available upon request.

that will be implemented as the optimal solution of the *IEA* (i.e. θ^*), according to the SDC-rule (see Sect. 3.1).

In this respect, an *asymmetric* influence of climate risks appears to be present: the poor country always establishes the *IEA*'s target when it faces a low risk (a_L^S), irrespective of the global conditions and the climate risk in the North. A possible interpretation might lie in the fact that the poor country prefers to boost polluting economic growth when the climate risk is low. Indeed, in the poorest regions, the absolute difference between polluting and green profits and the limited economic losses due to local damage (i.e. L_i), make the green strategy less convenient once the abatement target is fixed.

By contrast, when S faces high climate risk (a_S^H), the outcomes depend on N . If the richer country has low environmental concerns (a_N^L), it will establish the *IEA* at a lower level because it faces higher opportunity costs. However, under the same high climatic risk ($a_S^H = a_N^H = 0.9$), S is again the bottleneck because it prefers to pursue economic growth and to put the burden of emission reduction on N . Indeed, given the historical economic inequality (i.e. $\Pi_N \gg \Pi_S$), N will curb more carbon emissions, in absolute terms, once the same abatement target in the *IEA* is established. Notably, extreme climate risks ($a_N^H - a_S^H$) do not necessarily lead to drastic mitigation interventions. In the case of high inequality and economic growth (*Case 4*), the *IEA* target will prove lower than the other cases. The trade-off between uncertain environmental damage and economic interests, together with the asymmetric historical responsibility, thus generates an *impasse*, as often observed in real *IEAs*.²¹

The abatement targets will prove more stringent in the case of higher economic growth (*Case 2 and 4*) independently of both renewable energy profitability (i.e. higher green profits π_E), and the polluting/green profits gap, suggesting that economic growth brings about more efforts to avoid detrimental environmental effects. Interestingly, in some cases the proposed abatement target is particularly ambitious ($\theta^* > 0.7$) although not unrealistic. Indeed, the last report drawn up by the IPCC (2018) suggested that, in order to avoid a temperature rise greater than 1.5 Celsius degree, it is necessary to reach net zero carbon emissions by 2050.

5.2 Evolutionary equilibrium and actual emissions

Table 3 reports the results of Game 2 and the GHG emissions emerging from the evolutionary equilibrium (as described in the green and orange panels in Fig. 1). A detailed explanation of the different *regimes* and the stability of the equilibria is exposed in "Appendix" B.3. Note that the overall GHG is computed as the sum of country emissions given the share of polluting firms in equilibrium, namely:

$$GHG = (1 - \beta_N^*) \cdot \Pi_N^P + (1 - \beta_S^*) \cdot \Pi_S^P.$$

Note that, in the above scenarios, the fraction of green firms in (evolutionary stable) equilibrium (β^*) does not vary across countries because the value of the parameters (c , u , δ and θ), that determine β^* is the same (see Eq. B.2 and B.4 in "Appendix" B.1). However, the fraction of green consumers (α^*) that makes firms indifferent is larger in the richer country (i.e. $\alpha_N^* > \alpha_S^*$) because they are needed more green consumers are required to offset

²¹ For instance, in the case of the Kyoto Protocol compliance with the treaty was not mandatory for many developing countries (non-Annex I Parties).

Table 1 The setting of the parameters of the numerical simulations

| North | | South | | $\Pi_N + \Pi_S$ | $\frac{\Pi_N}{\Pi_S}$ | Case | Climate risk ($i=\{N, S\}$) |
|-----------|-----------|-----------|-----------|-----------------|-----------------------|------------------------------------|---|
| π_N^E | π_N^P | π_S^E | π_S^P | | | | |
| 0.2 | 0.4 | 0.1 | 0.2 | 0.9 | 2 | (1) Low inequality Low growth | Low : $a_i^L = 0.1$ High : $a_i^H = 0.9$ |
| 0.2 | 0.8 | 0.1 | 0.4 | 1.5 | 2 | (2) Low inequality High growth | Cons. Environ. Attitude |
| 0.3 | 0.7 | 0.1 | 0.2 | 1.3 | 3.3 | (3) igh inequality Low growth | Low: $c_L = 0.3$; $\delta_L = 0.33$ |
| 0.6 | 0.9 | 0.1 | 0.4 | 2.0 | 3 | (4) High inequality High growth | High: $c_H = 0.7$; $\delta_H = 0.9$ |

The right-hand matrix reports the values of the parameters regarding green (π^E) and brown (π^P) profits in both countries (N, S). Depending on the level of economic development ($\Pi_N + \Pi_S$) and cross-country technological inequality (Π_N/Π_S), four alternative cases are identified (see last column). The left-hand matrix reports, on the top, the distribution of climate risks, i.e. either low (a^L) or high (a^H). The bottom part shows the values—either low or high—of the parameters associated to consumer willingness to finance green start-ups (c) and environmental attitude (δ). Note that the opportunity costs and the consumers' utility are the same across all the simulation (i.e. $u = 0.1, \gamma = 1, b = 0.5$ and $d = 0.5$).

Table 2 Optimal bargaining solution from IEA in Game 1 (i.e. θ^*)

| | | | | | |
|----------------|-------------------|-------------------|----------------|-------------------|-------------------|
| Case 1) | a_N^L | a_N^H | Case 2) | a_N^L | a_N^H |
| a_S^L | 0.11 ^S | 0.00 ^S | a_S^L | 0.33 ^S | 0.19 ^S |
| a_S^H | 0.46 ^N | 0.73 ^S | a_S^H | 0.59 ^N | 0.80 ^S |
| Case 3) | a_N^L | a_N^H | Case 4) | a_N^L | a_N^H |
| a_S^L | 0.13 ^S | 0.00 ^S | a_S^L | 0.26 ^S | 0.15 ^S |
| a_S^H | 0.64 ^N | 0.80 ^S | a_S^H | 0.85 ^S | 0.59 ^S |

Each entry of each matrix represents the solution of Game 1 under the four alternative cases and climatic risk distribution (i.e. a^H high or a^L low) defined in Table 1. The superscripts ^S (South) and ^N (North), associated to each number, stand for the country that acts as the bottleneck in the IEA. Recall that the possible values of θ^* range from 0 (no abatement) to 1 (zero emissions)

the absolute difference between polluting and green profits. This gap is higher in absolute terms in the North in all four cases.

When θ^* is enforced by law in each country, the economic structure will determine the actual outcome. Here, what is crucial is the role played by the consumers' environmental attitude (δ). When it is low (δ_L), the actual environmental outcome is always worse than what is agreed in the IEA, irrespective of any free-riding behaviour. In some cases, the IEA is utterly ineffective ($\beta^* = 0$) due also to the unfavourable consumers' environmental attitude. This is a noteworthy result that shows the necessity of bottom-up initiatives for environmental policies to succeed. It also shows another source of failure of IEAs other than free-riding that depends on the actual socio-economic structure.

The only exception is given by the case of extreme climate risks that generate such an environmental target that the pay-off structure of consumers and firms is greatly affected. Interestingly, in these cases our model captures the over-performance

Table 3 Results Game 2

| Evolutionary stable share of green firms (β^*) | | | | | | | | | | | |
|--|------|---------|---------|---------|------|--|---------|---------|---------|---------|---------|
| Low Environmental Attitude (δ_L) | | | | | | High Environmental Attitude (δ_H) | | | | | |
| Case 1) | | a_L^N | a_H^N | Case 2) | | a_L^N | a_H^N | Case 1) | | a_L^N | a_H^N |
| a_L^S | 0.00 | 0.00 | a_L^S | 0.03 | 0.00 | a_L^S | 0.00 | 0.00 | a_L^S | 0.23 | 0.08 |
| a_H^S | 0.10 | 1.00 | a_H^S | 0.31 | 1.00 | a_H^S | 0.10 | 1.00 | a_H^S | 1.00 | 1.00 |
| Case 3) | | a_L^N | a_H^N | Case 4) | | a_L^N | a_H^N | Case 3) | | a_L^N | a_H^N |
| a_L^S | 0.00 | 0.00 | a_L^S | 0.00 | 0.00 | a_L^S | 0.02 | 0.00 | a_L^S | 0.15 | 0.04 |
| a_H^S | 0.38 | 1.00 | a_H^S | 1.00 | 0.31 | a_H^S | 1.00 | 1.00 | a_H^S | 1.00 | 1.00 |

| Actual GHG emission | | | | | | | | | | | |
|---|------|---------|---------|---------|------|--|---------|---------|---------|---------|---------|
| Low Environmental Attitude (δ_L) | | | | | | High Environmental Attitude (δ_H) | | | | | |
| Case 1) | | a_L^N | a_H^N | Case 2) | | a_L^N | a_H^N | Case 1) | | a_L^N | a_H^N |
| a_L^S | 0.60 | 0.60 | a_L^S | 1.17 | 1.20 | a_L^S | 0.60 | 0.60 | a_L^S | 0.93 | 1.11 |
| a_H^S | 0.54 | 0.00 | a_H^S | 0.83 | 0.00 | a_H^S | 0.18 | 0.00 | a_H^S | 0.00 | 0.00 |
| Case 3) | | a_L^N | a_H^N | Case 4) | | a_L^N | a_H^N | Case 3) | | a_L^N | a_H^N |
| a_L^S | 0.90 | 0.90 | a_L^S | 1.30 | 1.30 | a_L^S | 0.88 | 0.90 | a_L^S | 1.11 | 1.24 |
| a_H^S | 0.56 | 0.00 | a_H^S | 0.00 | 0.90 | a_H^S | 0.00 | 0.00 | a_H^S | 0.00 | 0.00 |

Each entry in the top part represents the evolutionary stable share of green firms (β^*) under the four alternative cases and climatic risks distribution (i.e. a^H high or a^L low) defined in Table 1. Each entry in the bottom part represents the actual global carbon pollution (GHG)—as the sum of emissions from each country—under the four alternative cases and climatic risks distribution (i.e. a^H high or a^L low) defined in Table 1, while GHG is the global emissions as the sum of air pollution from each country

observed in the real world, i.e. the country attains a greater abatement in emissions than what was ratified. Thus, when the top-down intervention is particularly strong the individual environmental attitude does not play a significant role. By contrast, when the environmental attitude is high the system converges towards a green-green equilibrium even with weaker targets (θ). This confirms previous findings according to which local participation can have the positive impact of accelerating the process of cleaning up production and saving resources for alternative uses, because they allow governments to fix less stringent standards and avoid additional expenditures on restoring environmental damage (Coenen 2009).

Our model also clarifies the double-edge role played by economic growth. On the one hand, when it is higher the abatement target tends to be more stringent. On the other hand, overall GHG emissions are greater even when the share of green firms is relatively high. This underlines the issue of the scale of the economy and suggests that *size matters*. Thus, the feasibility of low-carbon transition plans might be weaker in the case of sustained economic growth, unless clean production becomes exceptionally profitable (i.e. $\pi_E > \pi_P$). This concern finds confirmation in the lack of evidence for alleged GDP-CO₂ decoupling (Aslanidis and Iranzo 2009; Parrique et al. 2019) that might be due to the high cost of dismantling existing structures and to the fact that renewable energies are still not sufficient to cover the increasing global energy demand.

6 Discussion

This study analysed under what conditions international environmental agreements (*IEAs*) can be effective in terms of fulfilment of the ratified emission abatement targets. To this purpose, we considered two asymmetric countries—with different economic structures and technological development—negotiating emission reductions and then compared their actual environmental performance. It emerged that country-specific socio-economic dynamics are crucial in understanding why the *IEAs* might (or might not) fail. This claim was supported by the use of an analytical apparatus together with numerical simulations.

From a theoretical point of view, we offered an innovative perspective on which to ground multi-scale analysis. The main contribution with respect to the current literature on *IEAs* lies in modelling the country-specific economic structure which gives further insights to explain the gap between the promises of agreements and actual results. To the best of our knowledge, this study is the first to propose a *two-nested-game* approach where the outcomes of two different games are tied in a consistent framework. The proposed methodology might open the door to further developments in this field and shed light on how to bridge the micro- and macro- scales.

Although simplifications are essential, the effects of simplifying assumptions should not be ignored when interpreting the results (Madani 2013). In this respect, we acknowledge that there exist other drivers linked to climate change that might be relevant, such as: population growth, stock of emitted greenhouse gases remaining in the atmosphere, international trade, and time and costs required for conversion from a fossil-based to a zero-carbon economy. Future research might benefit from the inclusion of such factors, although the inclusion of too many variables, in a single tractable model, represents a mathematical and computational challenge.

Several implications may be drawn from our findings. First, contrary to what is assumed in the literature, our model shows that “global solutions”, negotiated at the international level, do not automatically produce the expected effects. Our framework allows a more accurate picture to be defined—including a variety of real-case results that would have been hidden in a standard context—in which even country-specific environmental over-performance (i.e. when carbon emissions are lower than what ratified) might emerge. Moreover, the assumption of *free-riding* might be misleading when dealing with collective entities and it is not necessary to explain the failure of the *IEAs*. Rather, at the country scale, environmental performance represents the outcome of a complex system resulting from the interplay of a large population of economic agents (consumers and firms). Indeed, our results suggest that the same *IEA* target may lead to diverging results depending on the country-specific socio-economic structure without assuming any free-riding behaviour.

Second, alternative combinations of climate risks distribution, economic growth, and cross-country technological inequalities generate non-trivial consequences. The impact of potential climatic economic losses generates asymmetric responses: the poorer countries prefer to fix lower abatement targets, all else kept constant. However, cross-country inequality appears to play a role when the climate risks are extreme. Contrary to what might be expected, the abatement target, ratified in the *IEA*, is lower because the poor country decides to place the burden on the richer one. This might explain why, during the Kyoto Protocol in 1997, compliance with the treaty was not mandatory for many developing countries (non-Annex I Parties). Moreover, “size matters”. Higher economic

growth, albeit translating into higher international abatement targets, results in higher GHG emission, thus confirming the scepticism regarding the possibility of achieving absolute decoupling between GDP and emissions.

Third, our approach highlighted the fact that the relation between justice and environmental targets is crucial to attain a low-carbon transition, as recently acknowledged in the literature (D'Alessandro et al. 2020). Increasing inequality might lead to adverse environmental attitudes (i.e. δ low) that might hamper the process of low-carbon transition implemented via energy policies (e.g. as shown by the protests of the “Gilet Jaune” in France). These observations entail that governments should not simply impose environmental standards by law. Rather, top-down policies and bottom-up interventions should be coordinated; otherwise, they might fail if undertaken in isolation.

Appendix A: Analytical analyses of the IEA (θ)

Appendix A.1: Business-as-Usual (BAU)

In the absence of climate change damage (i.e. a_i is null), each country maximizes its own welfare independently. In this case, each country i chooses the business-as-usual (BAU) solution (θ_i^{BAU}) so that it maximizes Π_i with respect to θ_i . The first-order condition ($\frac{\partial \Pi_i}{\partial \theta_i} = 0$) returns the optimum level of emissions under the business-as-usual hypothesis:

$$\theta_i^{BAU} = 1 - d_i \frac{(\pi_{P,i} - \pi_{E,i})}{\pi_{P,i}^2}. \tag{A.1}$$

Note that $0 \leq \theta_i^{BAU} \leq 1$ always²² and that, given d_i and $\pi_{E,i}$, it increases with respect to $\pi_{P,i}$ because the government must fix more stringent environmental standards in order to compensate the local ecological damages. On the other hand, a rich country—with $\pi_{P,i}$ high—might prefer to fix stringent environmental standards if it has an advanced green technology, and then when $\pi_{E,i}$ is high as well (i.e. $\pi_{P,i} - \pi_{E,i} \simeq 0$). Indeed, in this case, it would be an advantage to speed up the green transition because it can yield high profits without hurting the environment, so avoiding public expenditure to recover possible environmental damages.

Figure 2 shows the distribution of national environmental targets in case of no IEA (i.e. $a = 0$) depending on the combination of green and brown profits, that range from low (0.1) to high (0.9). First, when $\pi_E > \pi_P$ the country have convenience to set the maximum environmental target (i.e. $\theta^{BAU} = 1$, 100% of abatement) because it can achieve economic growth without detrimental environmental effects. Second, the target decreases as the opportunity cost (d) increase because the economic loss from lower emissions, due to lower production of green firms, offset the environmental benefits. These two results are particularly evident in the bottom-right matrix, where θ is particularly low (0.2) when the difference between polluting and green profits is maximum.

²² Since θ_i^{BAU} might take negative values, when $\pi_{P,i}$ decreases, we bound it to be non-negative, to say it is null when the country opts for no environmental laws.

Appendix A.2: Static comparative Analysis

Case I—different climatic risks let us assume that both countries have the same marginal benefits, $b_N = b_S = b$ and $d_N = d_S = d$, but different climatic risks: $a_N = \tilde{z} \cdot a_S$, with $\tilde{z} > 0$. Differences in a might be due to different weights put to the environment—which can be related to the economic development of a region—or to the geographical location (e.g. Italy may suffer more from sea level rise than Russia). Let define $\lambda_i = \frac{b_i}{a_i} > 0$ as the *benefit–risk ratio* of country i , given by the marginal benefit of local production and potential losses from global emissions..

Proposition 1 *The possibility that country S will be the bottleneck in the non-coordinated game (i.e. if $i = S$ and $\theta_N > \theta_S$, then it results that $\theta^{*NC} = \theta_S$), under different climatic risks, increases with technological inequality.*

Proof To prove it we have to find the conditions, with respect to \tilde{z} , such that $\Theta_N > \Theta_S$. Let us substitute Eq. (6) and Eq. (7) in the above-stated relations: $b_N = b_S = b$, $d_N = d_S = d$, $\lambda_S = \frac{b_S}{a_S} > 0$, $\pi_{P,N} = m \cdot \pi_{P,S}$, $\pi_{E,N} = m \cdot \pi_{E,S}$, and $\pi_E = n \cdot \pi_P$. Then, $\Theta_N > \Theta_S$ holds true if and only if:

$$\tilde{z} > \bar{z} \equiv 1 + \lambda_S \cdot \frac{(1 - m)}{(1 + m)} \tag{A.2}$$

Recalling that m is a measure of cross-country inequality, that is minimum when $m = 1$ (i.e. the two countries have the same technological and income level); hence, to demonstrate that increasing inequality reduces the threshold, we compute the following limits:

$$\lim_{m \rightarrow 1} \bar{z} = 1 \tag{A.3}$$

$$\lim_{m \rightarrow +\infty} \bar{z} = 1 - \lambda_S \tag{A.4}$$

Since $\lambda_S > 0$, then the threshold \bar{z} is always lower in case of greater inequality. Moreover, since the marginal benefit is the same, the threshold is smaller when the climate risk of country S is low. □

This proposition shows that historical inequality (m) and uneven distribution of climate risks have effects on current environmental decisions, inducing countries to agree to lower emission abatement levels (Tavoni et al. 2015). As seen, in case of high cross-country inequality, if the climatic risk of the poorer country is relatively low it will propose a low environmental target that will become the Nash equilibrium as S is the bottleneck of the IEA. An example is given by China which is likely to suffer from extreme climate events (e.g. desertification) but it prefers to further develop the industrial production and it is thus less concerned about emission reductions.

Case II—different opportunity costs let us now consider the opposite case of equal climate damage, $a_N = a_S = a$ but different opportunity costs of abatement: $b_N = \hat{z} \cdot b_S$, with $\hat{z} > 0$.

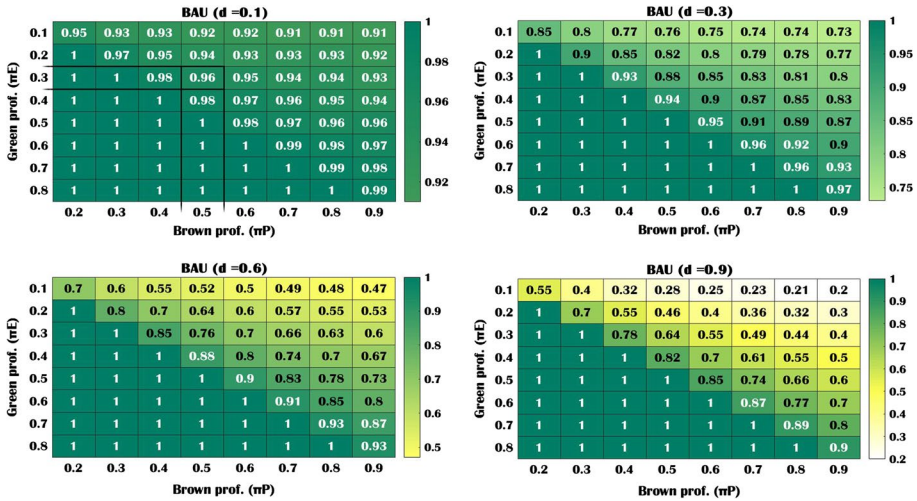


Fig. 2 National environmental law. Distribution of environmental targets, in case of no IEA, depending on the combination of green (π_E) and polluting (π_P) profits and opportunity cost (d)

Proposition 2 *The possibility that country S will be the bottleneck in the non-coordinated game (i.e. if $i = S$ and $\theta_N > \theta_S$, then it results that $\theta^{*NC} = \theta_S$), under different opportunity costs, depends on the benefit-to-risk ratio of S .*

Proof In this case we repeat the same procedure of Case I, but considering that $a_N = a_S = a$ and $b_N = \hat{z} \cdot b_S$. Then, $\Theta_N > \Theta_S$ holds true if and only if:

$$\hat{z} < \bar{z} \equiv \frac{a \cdot (1 + m)}{a(1 + m) + b_S \cdot (1 - m)} \tag{A.5}$$

Since the threshold is non-linear in λ_S , we compute the following limits to analyse the different cases:

$$\bar{z}_1 = \lim_{m \rightarrow 1} \bar{z} = 1; \quad \bar{z}_\infty = \lim_{m \rightarrow +\infty} \bar{z} = \frac{1}{1 - \lambda_S}. \tag{A.6}$$

□

Appendix B: Analytical discussion of evolutionary equilibria

In what follows, we provide a *ceteris paribus* analysis of the different evolutionary stable states, of both households and firms' share of green strategies (α, β),²³ depending on the stringency of the international environmental law (θ). The solutions come from the

²³ We recall that the proportion of polluting agents is the complement of the green one, since the two shares sum up to 1.

replicator dynamics (see Eq. (8) and (9)): in case of the households the share of green consumers stabilizes when $H_E = H_P$ that ensures that $\dot{\alpha} = 0$, while in case of firms when $F_E = F_P$ that ensures that $\dot{\beta} = 0$. The discussion of the different fixed points proceeds for households and firms separately because—given θ —it is possible to establish analytically the share of firms only once the proportion of households stabilizes, and vice versa. Hence, in the households section we will discuss the different evolutionary stable states of green firms given the proportion of households, and vice versa.

For the sake of clarity, we recall the meaning of the different symbols:

- u : constant level of utility from consumption;
- δ and c : environmental attitude and extra-cost to finance green start-ups, respectively;
- π_P and π_E : profits from polluting and green production, respectively;
- γ : multiplicative factor of environmental damages.

Appendix B.1: Households

Proposition 3 *Households choose the green (polluting) strategy if and only if $H_E > H_P$ ($H_E < H_P$).*

Proof The demonstration is rather simple and relies on the common assumption of individual as maximizer of utility. Thus, if the expected value of a strategy returns an higher pay-off, it will be always preferred.

In particular we have the following cases:

- if $\beta = 0$, then $H_E > H_P$ if and only if $\theta > \theta_0 \equiv \frac{u}{\delta}$. Note that when $\beta = 0$ so does H_E ; therefore, the only way to discourage polluting consumption is to fix θ high enough to make H_P (temporarily) negative;²⁴
- if $\beta = 1$, then $H_E > H_P$ always.

□

In case $0 < \theta \leq \theta_0 \equiv \frac{u}{\delta}$, the solutions are given either by

$$\beta_0^* \equiv \frac{1}{2}, \text{ if } 0 < \theta < \min\{\theta_0, \frac{1}{2}\}, \tag{B.1}$$

or by

$$\beta_1^* \equiv \frac{\theta(c + \delta) + \delta + \sqrt{\Delta_\beta}}{2(c + \delta)} \text{ if } \frac{1}{2} < \theta < \theta_0, \tag{B.2}$$

where $\Delta_\beta = [\theta(c + \delta) + \delta - 2u]^2 - 4(c + \delta)(\delta\theta - u)$. From the last term of Δ_β , it is straightforward that $\theta < \theta_0$ is a sufficient condition for $\Delta_\beta > 0$. It holds that if $\delta \leq \frac{c^2 + 4u^2}{4u}$ the determinant is always positive, otherwise

²⁴ We might avoid negative utilities simply by assuming any scale factor—out of the diagonal—big enough to compensate the gap. However, this issue does not alter the nature of the game.

$$\Delta_\beta \geq 0 \iff \theta \leq \theta_1 \equiv \frac{\delta + 2u - 2\sqrt{(\delta - c)u}}{c + \delta}. \tag{B.3}$$

If the environmental attitude is sufficiently low (i.e. $\delta < u$), there is a single intersection between H_E and H_P for any value of $\theta \in [0, 1]$. In case $\beta < \min\{\beta_0^*, \beta_1^*\}$, the expected reward of the polluting strategy is greater than the green one, while if $\beta > \max\{\beta_0^*, \beta_1^*\}$ the reverse holds. When $\theta > \theta_0$, the two curves— H_E and H_P —can be either secant ($\Delta_\beta > 0$), tangent ($\Delta_\beta = 0$), or without any point in common ($\Delta_\beta < 0$). While, in case $\theta_0 < \theta \leq \theta_1$ ($= \delta \leq \frac{c^2 + 4u^2}{4u}$), H_E and H_P have two intersections (see Fig. 3b): the first one is β_0^* if $\theta_0 < \theta < \min\{1/2, \theta_1\}$, or β_1^* if $\max\{\theta_0, 1/2\} < \theta < \min\{1, \theta_1\}$. The second solution (β_2^*) is given by:

$$\beta_2^* = \frac{\theta(c + \delta) + \delta - \sqrt{\Delta_\beta}}{2(c + \delta)}. \tag{B.4}$$

In this case, if $0 \leq \beta < \beta_2^*$ and $\max\{\beta_0^*, \beta_1^*\} < \beta \leq 1$, then the expected pay-off of the green strategy is greater than that of the polluting one, while for $\beta_2^* < \beta < \beta_{0,1}^*$ the reverse holds. Figure 3c shows a case in which there is no intersection between the two expected pay-off, to say the green strategy is always preferred for any value of β .

Figure 3 shows the possible equilibrium deriving from the expected households' pay-off as a function of the share of green firms (β). This might clarify the interpretation of the numerical simulations in Section 5.

Appendix B.2: Firms

Proposition 4 *Knowing the share of households which choose the green or the polluting strategy in the previous period, firms prefer the green (polluting) production if and only if the expected pay-off of \mathcal{F}_E is greater (lower) than the expected pay-off of \mathcal{F}_P .*

Proof Let us define, neglecting for the time specification, $F_E = E(\mathcal{F}_E) = \alpha f_E$ and $F_P = E(\mathcal{F}_P) = (1 - \alpha)f_P$ the expected pay-off of the green and the polluting production, respectively, then the expected firms' pay-off depend on both α and β . Note that F_E is an increasing function of α , such that: $F_E = 0$ when $\alpha = 0$ and $F_E > 0$ when $\alpha = 1$.²⁵ On the other hand, the expected pay-off of the polluting strategy is decreasing in α if $\beta > \theta$. Instead, when $\beta < \theta$ the slope of the function F_P depends on the sign of $\pi_P - \frac{\gamma(\theta - \beta)}{1 - \beta}$ which expresses the difference between the whole profits of polluting industries and the (monetary evaluation) of the environmental damages. Obviously, when this last expression is negative, that is when $\beta < \tilde{\beta} \equiv \frac{\gamma\theta - \pi_P}{\gamma - \pi_P}$, then F_E is greater than F_P for any value of α . \square

Figure 4 shows the resulting interceptions between F_E and F_P in the plane $\{\theta, \beta\}$. When $\beta > \theta$ there is an interior value of α (α_0^*), such that firms are indifferent between the two strategies, which does not depend on β . When instead $\tilde{\beta} \equiv \frac{\gamma\theta - \pi_P}{\gamma - \pi_P} < \beta < \theta$, there is an interior value of α (α_1^*), such that $F_E = F_P$, but this value is an increasing function of β . Note

²⁵ The value of F_E at $\alpha = 1$ depends on the relation between θ and β . If $\beta < \theta$ then $F_E = \pi_E + \frac{c(\theta - \beta)}{\beta}$, otherwise $F_E = \pi_E$.

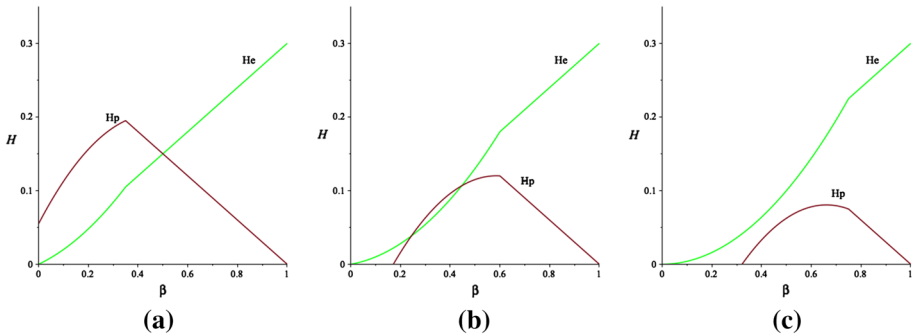


Fig. 3 Analysis of households equilibrium. Green and red lines are the expected pay-off of green and polluting consumers, respectively. Panel (a) shows the case of single intersection, panel (b) the case of two interior equilibrium (and one evolutionary stable state), while panel (c) the case of no intersection and no interior equilibrium

that $\theta > \frac{\pi_p}{\gamma}$ is a necessary condition to induce at least one firm to deviate from the polluting convention (i.e. when $\alpha = 0$ and $\beta = 0$). More precisely, the two possible solutions are:

$$\alpha_0^* = \frac{\pi_p}{\pi_p + \pi_E}, \tag{B.5}$$

$$\alpha_1^* = \frac{\beta[\gamma(\theta - \beta) - (\pi_p + \pi_E)(1 - \beta)] + \sqrt{\Delta_\alpha}}{2c[(\theta - \beta)(1 - \beta)]} \tag{B.6}$$

where Δ_α is always positive.²⁶

These results, combined with those of the previous subsection, determine analytically the equilibria of the evolutionary game.

Appendix B.3: Regimes

Figure 5 shows all the possible outcomes of the evolutionary game, depending on households' and firms' pay-off. Households and firms change their behaviour according to the replicator dynamics described by Eq. (8) and (9). Depending on the value of θ five *Regimes* (R_r) emerge, showing the qualitative change of the dynamic properties of the system.²⁷ For the sake of clearness, we assume that the initial condition of the economy is in the polluting convention where $\beta = \alpha = 0$ and that the government establishes a certain level of environmental standard ($\theta > 0$). Figure 5 shows the phase diagram for each Regime. Note that the black circles indicate the stationary stable equilibria and the light-green horizontal line is the environmental target ratified in the *IEA* (θ).

²⁶ Note that the other solution in α of $F_E = F_p$ is always negative. Moreover, $\Delta_\alpha = \beta^2\{(1 - \beta)[2\gamma(\pi_p + \pi_E)(\beta - \theta) + (\pi_p + \pi_E)(1 - \beta)] + 4c(\pi_p - \gamma)(2\theta + \beta^2) + \gamma^2(\theta - \beta)^2\} + 4c\{\beta\pi_p(\beta^2(2 + \theta) - (\theta + \beta)) + \gamma(\beta^2\theta^2 - \beta - 2\theta) - \theta^2\}$

²⁷ Value of parameters in Fig. 5: (a) $\delta = 0.7, c = 0.3, u = 0.35, \gamma = 1.1, \pi_E = 0.1, \pi_p = 0.3, \theta = 0.15$; (b) $\delta = 0.5, c = 0.5, u = 0.35, \gamma = 1.5, \pi_E = 0.1, \pi_p = 0.3, \theta = 0.6$; (c) $\delta = 0.5, c = 0.1, u = 0.15, \gamma = 1.5, \pi_E = 0.1, \pi_p = 0.3, \theta = 0.25$; (d) $\delta = 0.5, c = 0.1, u = 0.15, \gamma = 1.5, \pi_E = 0.1, \pi_p = 0.6, \theta = 0.35$; (e) $\delta = 0.5, c = 0.1, u = 0.15, \gamma = 1.5, \pi_E = 0.1, \pi_p = 0.6, \theta = 0.45$; (f) $\delta = 0.5, c = 0.1, u = 0.15, \gamma = 1.5, \pi_E = 0.1, \pi_p = 0.6, \theta = 0.55$.

Fig. 4 Graphical analysis of firms' space of solutions. The intersection between the expected pay-off of green and polluting strategies in the plane $\{\theta, \beta\}$

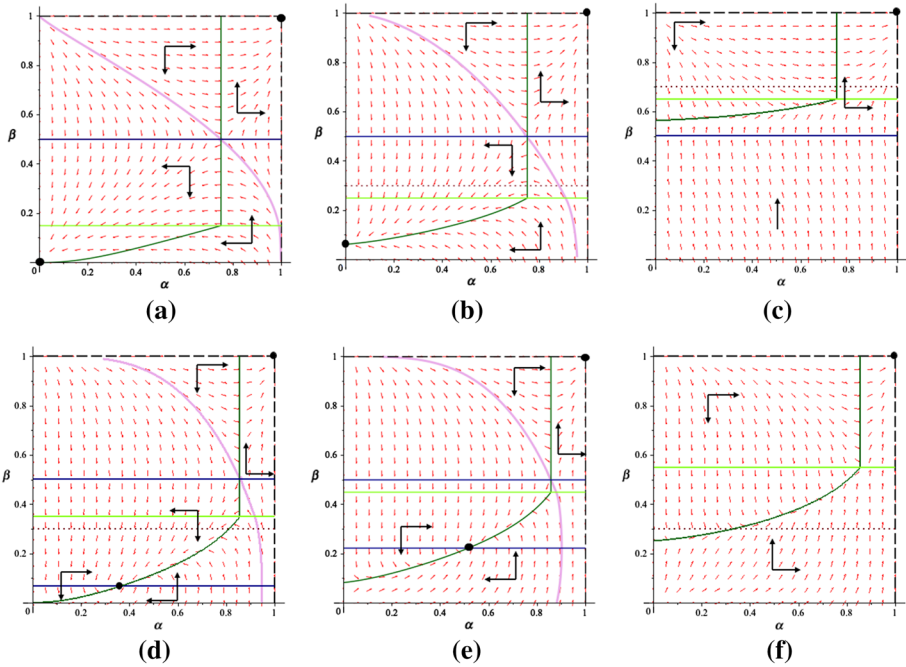
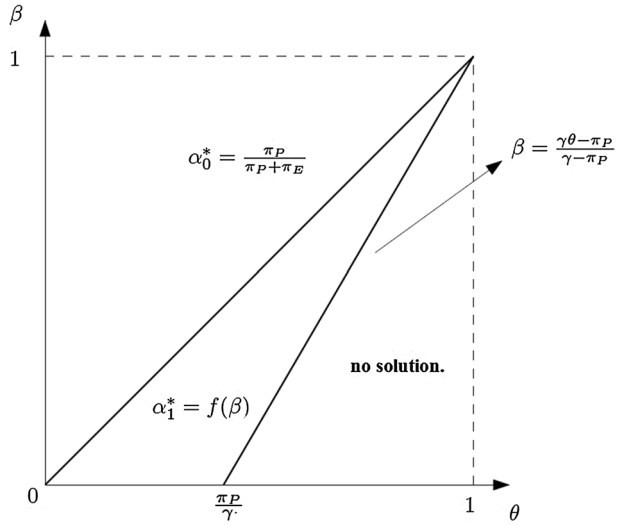


Fig. 5 Phase diagram of the dynamical system. The red arrows show the directions of the trajectory. The isocline of the share of green firms (β) is given by the dark green curve. The isocline of the share of green households (α) is given by the horizontal dark blue line(s). The value of θ is the light green horizontal line. The dot horizontal line is the value of u/δ . The magenta curve shows the basin of attraction of the two conventions, if applicable

R_1 : when $0 \leq \theta < \min\{\frac{u}{\delta}, \frac{\pi_p}{\gamma}\}$, the isoclines and the phase diagram of the system are shown in panel. 5(a). In this case there are no interior (locally) fixed points. The introduction of an environmental law ($\theta > 0$) is not sufficient to induce the system to detach from the polluting productive convention. The possible explanations of the failure of the policy (θ) can be found in:

- consumers are not enough aware of the potential environmental damages (δ low);
- high level of utility from material consumption (u high);
- high gap between green and dirty production ($\pi_p - \pi_E$ high) that makes convenient for the polluting firms to pay for the environmental damages instead of converting their production toward renewable energies.²⁸

R_2 includes two cases: R_{2a} is observed when $\frac{\pi_p}{\gamma} < \theta < \frac{u}{\delta}$, the polluting convention becomes unstable because the environmental law is high enough to induce the start-up of new green firms as long as $\beta < \bar{\beta}$, that is the condition for which $F_E > F_P$. If $\bar{\beta} < \min\{\beta_0^*, \beta_1^*\}$, then all the trajectories departing from the polluting convention converge to the fixed point with coordinates $(\alpha^*, \beta^*) = (0, \bar{\beta})$. This is a corner solution (Fig. 5b) that signals the imbalance between the high cost for polluting firms and the low attitude of households to environmental concerns.

R_{2b} If $\gamma < \pi_p$ and $\bar{\beta} > \max\{\beta_0^*, \beta_1^*\}$ the system detaches from the corner solution. As long as β increases, the households prefer to choose the green production. This process ends up when $(\alpha^*, \beta^*) = (1, 1)$. Thus, the only globally stable equilibrium is the green convention (see Fig. 5c).

R_3 When $\frac{u}{\delta} < \theta < \frac{\pi_p}{\gamma}$, households prefer to choose the green strategy so that they induce the firms to supply more green goods and services. The dynamical system is characterized by two locally stable fixed point, an interior point ($\alpha^* > 0, \beta^* > 0$) and the clean convention $(\alpha^*, \beta^*) = (1, 1)$. In this case, all the trajectories departing from the polluting convention join the interior equilibrium where $\beta^* \leq \theta$ (see Fig. 5d). The environmental policy has only a partial effect because θ is not high enough to induce the expected share of firms to shift their production from the polluting convention. Its impact is indirect and simply stands on the stimulus from the demand-side.

R_4 When $\max\{\frac{u}{\delta}, \frac{\pi_p}{\gamma}\} \leq \theta < \min\{\theta_1, 1\}$ households and firms prefer to choose the green strategy. Note that θ_1 is the threshold that ensures the stability of the interior equilibrium (see Eq. (B.3) in “Appendix” A.2.1 for the mathematical demonstration). The dynamical system defines two locally stable fixed points, the interior one and the green convention. In this regime, all the trajectories departing from the polluting convention end up to the interior equilibrium where $\beta^* \leq \theta$ (see Fig. 5e), accordingly, the same considerations of the previous regime hold true here.

R_5 When $\theta_1 < \theta < 1$ the only globally stable equilibrium is the green convention (see Fig. 5f) because the environmental law is sufficiently high to induce every agent to prefer the ecological strategy. Note that, in this case, the environmental target has not to be necessarily high to generate the low-carbon transition; rather, its success is strictly tied with the economic structure of the country and with the level of citizens’ environmental attitude.

Acknowledgements This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

²⁸ Note that given θ an increase in δ or γ may induce the system to depart from R_1 and to follow one of the other regimes.

Funding Open access funding provided by Università di Pisa within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Alvaredo, F., Chancel, L., Piketty, T., Saez, E., & Zucman, G. (2018). The elephant curve of global inequality and growth. *AEA Pap. Proc.*, *108*, 103–08.
- Anand, R. (2017). *International environmental justice: A North-South dimension*. Routledge.
- Antoci, A., Borghesi, S., Iannucci, G., & Ticci, E. (2019). Land use and pollution in a two-sector evolutionary model. *Struct. Change Econ. Dyn.*
- Aslanidis, N., & Iranzo, S. (2009). Environment and development: is there a Kuznets curve for CO2 emissions? *Appl. Econ.*, *41*(6), 803–810.
- Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxf. Econ. Pap.*, *46*, 878–894.
- Battaglia, M., Cerrini, E., & Annesi, N. (2018). Can environmental agreements represent an opportunity for green jobs? Evidence from two Italian experiences. *J. Clean. Prod.*, *175*, 257–266.
- Belloc, M., & Bowles, S. (2013). The persistence of inferior cultural-institutional conventions. *Am. Econ. Rev.*, *103*(3), 93–98.
- Bergset, L., & Fichter, K. (2015). Green start-ups—a new typology for sustainable entrepreneurship and innovation research. *J. Innov. Manage.*, *3*(3), 118–144.
- Bowles, S. (2009). *Microeconomics: behavior, institutions, and evolution*. Princeton: Princeton University Press.
- Cai, Y., Riezman, R., & Whalley, J. (2013). International trade and the negotiability of global climate change agreements. *Econ. Model.*, *33*, 421–427.
- Coenen, F. (2009). Public participation and better environmental decisions. *The promise and limits of participatory processes for the quality of environmentally related decision-making*, p. 209.
- Coronese, M., Lamperti, F., Keller, K., Chiaromonte, F., & Roventini, A. (2019). Evidence for sharp increase in the economic damages of extreme natural disasters. *PNAS*.
- D'Alessandro, S., Cieplinski, A., Distefano, T., & Dittmer, K. (2020). Feasible alternatives to green growth. *Nat. Sustain.*, *3*(4), 329–335.
- Endres, A., & Finus, M. (1999). International environmental agreements: How the policy instrument affects equilibrium emissions and welfare. *J. Inst. Theor. Econ. (JITE)/Zeitschrift für die gesamte Staatswissenschaft* *155*(3), 527–550.
- Finus, M. (2002). *Game theory and international environmental cooperation: any practical application?*, Chapter 2, pp. 9–104. Edward Elgar Cheltenham.
- Finus, M., Saiz, M. E., & Hendrix, E. M. (2009). An empirical test of new developments in coalition theory for the design of international environmental agreements. *Environ. Dev. Econ.*, *14*(1), 117–137.
- Gelves, A., & McGinty, M. (2016). International environmental agreements with consistent conjectures. *J. Environ. Econ. Manage.*, *78*, 67–84.
- Goeschl, T., & Perino, G. (2017). The climate policy hold-up: Green technologies, intellectual property rights, and the abatement incentives of international agreements. *Scand. J. Econ.*, *119*(3), 709–732.
- Grasso, M. (2011). The role of justice in the north-south conflict in climate change: the case of negotiations on the adaptation fund. *Int. Environ. Agreem. Polit. Law Econ.*, *11*(4), 361–377.

- Hannam, P. M., Vasconcelos, V. V., Levin, S. A., & Pacheco, J. M. (2017). Incomplete cooperation and co-benefits: deepening climate cooperation with a proliferation of small agreements. *Clim. Change*, *144*(1), 65–79.
- Hatase, M., & Matsubayashi, Y. (2019). Does government promote or hinder capital accumulation? Evidence from Japan's high-growth era. *Struct. Change Econ. Dyn.*, *49*, 245–265.
- IPCC (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 (First ed.). In Press. Masson-Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (eds.).
- Keohane, R. O., & Victor, D. G. (2011). The regime complex for climate change. *Perspect. Polit.*, *9*(1), 7–23.
- Madani, K. (2013). Modeling international climate change negotiations more responsibly: Can highly simplified game theory models provide reliable policy insights? *Ecol. Econ.*, *90*, 68–76.
- Marrouch, W., Chaudhuri, A. R., et al. (2016). International environmental agreements: Doomed to fail or destined to succeed? A review of the literature. *Int. Rev. Environ. Res. Econ.*, *9*(3–4), 245–319.
- Nasiritousi, N., Hjerpe, M., & Linnér, B.-O. (2016). The roles of non-state actors in climate change governance: understanding agency through governance profiles. *Int. Environ. Agree. Polit. Law Econ.*, *16*(1), 109–126.
- Olivier, J., Peters, J., & Janssens-Maenhout, G. (2013). *Trends in global CO2 emissions 2012 report*. ETDE: Technical report.
- Parrique, T., Barth, J., Briens, F., Kerschner, C., Kraus-Polk, A., Kuokkanen, A., & Spangenberg, J. (2019). *Decoupling debunked: Evidence and arguments against green growth as a sole strategy for sustainability*. Brussels, Belgium: European Environmental Bureau.
- Pavlova, Y., & de Zeeuw, A. (2013). Asymmetries in international environmental agreements. *Environ. Dev. Econ.*, *18*(1), 51–68.
- Ponticelli, J., & Voth, H.-J. (2020). Austerity and anarchy: Budget cuts and social unrest in Europe, 1919–2008. *J. Comp. Econ.*, *48*(1), 1–19.
- Richard, S. T. (1995). The damage costs of climate change toward more comprehensive calculations. *Environ. Resource Econ.*, *5*(4), 353–374.
- Roberts, J. T., & Parks, B. (2006). *A climate of injustice: Global inequality, north-south politics, and climate policy*. Cambridge: MIT press.
- Santos, F. C., & Pacheco, J. M. (2011). Risk of collective failure provides an escape from the tragedy of the commons. *Proc. Nat. Acad. Sci.*, *108*(26), 10421–10425.
- Shi, X., Dong, C., Zhang, C., & Zhang, X. (2019). Who should invest in clean technologies in a supply chain with competition? *J. Clean. Prod.*, *215*, 689–700.
- Sigmund, K. (2010). *The Calculus of Selfishness*. Princeton: Princeton University Press.
- Steg, L. (2016). Values, norms, and intrinsic motivation to act proenvironmentally. *Annu. Rev. Environ. Resour.*, *41*, 277–292.
- Tavoni, M., Kriegler, E., Riahi, K., Van Vuuren, D. P., Aboumahboub, T., Bowen, A., et al. (2015). Post-2020 climate agreements in the major economies assessed in the light of global models. *Nat. Clim. Change*, *5*(2), 119.
- Testa, F., Annunziata, E., Iraldo, F., & Frey, M. (2016). Drawbacks and opportunities of green public procurement: an effective tool for sustainable production. *J. Clean. Prod.*, *112*, 1893–1900.
- Turaga, R. M. R., Howarth, R. B., & Borsuk, M. E. (2010). Pro-environmental behavior: Rational choice meets moral motivation. *Ann. N. Y. Acad. Sci.*, *1185*(1), 211–224.
- Weibull, J. W. (1997). *Evolutionary game theory*. Cambridge: MIT press.
- Yang, Z. (2017). Likelihood of environmental coalitions and the number of coalition members: Evidences from an iam model. *Ann. Oper. Res.*, *255*(1–2), 9–28.
- Zhang, Y.-X., Chao, Q.-C., Zheng, Q.-H., & Huang, L. (2017). The withdrawal of the US from the Paris Agreement and its impact on global climate change governance. *Adv. Clim. Change Res.*, *8*(4), 213–219.