



John Rawls and compliance to climate change agreements: insights from a laboratory experiment

Klaudijo Klaser¹ · Lorenzo Sacconi^{2,3} · Marco Faillo¹

Accepted: 5 March 2021
© The Author(s) 2021

Abstract

The most evident shortcoming of the international agreements on climate actions is the compliance to their prescriptions. Can John Rawls's social contract theory help us to solve the problem? We apply the veil of ignorance decision-making setting in a sequential dictator game to study the compliance to climate change agreements and we test the model in a laboratory experiment. The veil of ignorance shows to be very powerful at inducing the subjects to converge on a sustainable intergenerational path. However, the voluntary compliance to the agreement still remains an open issue, because even small incentives to defect can undermine the compliance stability, and therefore break the whole sustainable dynamic.

Keywords International climate change agreements · Compliance · John Rawls · Social contract theory · Experimental economics

JEL Classification D63 · D64 · F64 · Q54

Society is indeed a contract [...] it becomes a partnership not only between those who are living, but between those who are living, those who are dead, and those who are to be born

Edmund Burke

✉ Klaudijo Klaser
klaudijo.klaser@unitn.it

Lorenzo Sacconi
lorenzo.sacconi@unimi.it

Marco Faillo
marco.faillo@unitn.it

¹ Dipartimento di Economia e Management, Università degli Studi di Trento, Via Inama 5, 38122 Trento, Italy

² Dipartimento di Diritto Pubblico Italiano e Sovranazionale, Università degli Studi di Milano, Via Festa del Perdono 7, 20122 Milano, Italy

³ EconomEtica, Centro interuniversitario per l'etica economica e la responsabilità sociale d'impresa, Università Milano Bicocca, Via Bicocca degli Arcimboldi, 8 - Edificio U7, 20126 Milano, Italy

[E]conomics needs to take account of the alternatives to utilitarianism that have been advanced in the past half century, such as the theory of justice of Rawls

Anthony Barnes Atkinson

1 Introduction

Climate change is a threat that looms over future generations but that is triggered by the consumption decisions taken in the present (Clements, 2015). Therefore, only acting beforehand the present generation can prevent the global warming and its negative consequences on future generations. At the same time, climate change is a global issue that cannot be tackled by means of the commitment of a minority of virtuous agents. Instead, problems like the global warming can be solved only through a certain degree of international co-operation (Stern et al., 2006): this is the main goal of the international agreements on climate actions.

However, “notwithstanding more than twenty years of international negotiations to establish limits, emissions of greenhouse gases continue to rise”¹ (Gardiner et al., 2016, p. 137), so much so that according to the World Meteorological Organization in 2017, the greenhouse gas levels in the atmosphere reached a new record (WMO, 2018). Thus, although mankind understood the dangers intrinsic in climate change long time ago (Nordhaus, 1993) and so far many international agreements have been formally stipulated between countries to contain harmful emissions, as a matter of fact, nations have proved to be incapable to limit the risk of damaging future generations. Two are the main cause of this constant failure of the international agreements on climate actions.

On the one hand, the success of the agreements concerning the reduction of greenhouse gases is essentially conditioned upon the distribution of costs between nations. Whatever is the contract, some states are expected to make grater sacrifices in limiting emissions, that is some nations are supposed to pay higher costs for the pursuit of the common goal (Brick & Visser, 2015; Gardiner, 2011a; Gardiner et al., 2016; Klinsky & Dowlatabadi, 2009; Lange et al., 2010; Sayegh, 2020). However, nobody is really available to pay more than the others. This forestalls any type of agreement because it is not clear which nations should bear the higher costs—those which produced more greenhouse gas in the past or those producing more emissions nowadays, for not considering the absolute vs. the per capita emissions criterion (Gardiner, 2011a).

On the other hand, even though so far more than a dozen of climate agreements have been formally signed—showing thus an ostensible shared willingness to overcome the obstacle mentioned above—compliance to such treaties has proved to be extremely fragile. Indeed, since reducing emissions are costly and since in the current geopolitical frame there are no global institutions which have the power to monitor or to sanction the defectors, single nations have a clear economic incentive to free ride (Brekke & Johansson-Stenman, 2008). Therefore, like in a standard voluntary contribution mechanism (Isaac & Walker, 1988), while it might be collectively rational to co-operate, from the individual point of view, it is rational to deviate from the agreement (Gardiner, 2011a; Lange et al., 2010).

¹ The first formal attempt to address climate change towards a common solution can be identified in the United Nations Framework Convention on Climate Change in 1992; the last one in the Paris Agreement of 2015. In the interval, we can count at least other 15 agreements to establish limits to greenhouse gas emissions.

However, without an international agreement aimed at reducing effectively greenhouse gas emissions, the world will soon head a tragedy of commons (Hardin, 1968; Ostrom et al., 1994) because the atmosphere, although big, can contain a limited quantity of greenhouse gases before these produce their harmful and irreversible effects on the whole mankind, and in particular on future generations. Starting from this general frame where the lack of compliance seems to be a widely observed behaviour, in this paper, we address some structural issues of climate change agreements through the social contract theory of John Rawls (1999). In particular, we focus on the voluntary compliance to the agreements.

John Rawls (1999) was the pioneer in extending the social contract model to the problem of the distribution of resources between generations. Rawls claims that if we prevent the present generation from knowing the specific moment of the history to which it belongs (veil of ignorance), the currently existing people can agree and voluntarily comply (sense of justice) with fair principles which regulate the consumption of resources between generations (Rawls, 1999, pp. 118–123). Thus, the veil of ignorance model helps (1) to mitigate the self-serving bias, which has been shown to play the major role in preventing effective agreements on greenhouse gas emissions (Kriss et al., 2011; Lange et al., 2010), and (2) to comply with the chosen rules in the absence of monitoring and sanctioning institutions. We applied and tested this model in a laboratory experiment.

The next Sections are organized as follows. Section 1 contextualizes the paper describing the standard approach to deal with distribution of resources between generations. Section 2 introduces the main elements of Rawls's social contract theory (Rawls, 1999), in particular, the idea of an agreement between contemporaries for intergenerational distributive principles. Section 3 recalls the previous attempts to test Rawls's intergenerational theory in a laboratory experiment. Section 4 describes our experimental design which captures the main features of the modern climate change agreements and it provides the predictive hypothesis. Section 5 reports the observed data. In Sect. 6, we comment the data. In the Conclusions, we summarize the main achievements, highlight the limits of our study and provide some insights for the future research.

2 The standard economic approach to the environmental issues

In the eighteenth century, Adam Smith (1976, 1994), when formulating his economic and moral theories about the free market and the impartial spectator, did not address any environmental issue: such as modelled, the invisible hand does not take into account the interests of future generations. Only after the industrialization process spread all over the world and Simon Kuznets's curve (1955) was reinterpreted under an environmental perspective (Dinda, 2004) the sustainability issue started to occupy an increasing share within the economic debate. Since then, many economists have tried to answer the following question: how can the current generation, in its own choices, take concretely into account the interests of future generations which have no voice at the round table—because not existing yet— but which are clearly direct stakeholders of the present decisions? Or in a simplified version, how can resources be distributed fairly across generations?²

² The answer to this question is make difficult by the many features that distinguish allocation of resources between generations (Meyer, 2016) from the standard redistributive issue between contemporaries (Lamont et al., 2016; Tremmel, 2009), “such as the non-identity problem [...], the co-operation between generations [...], motivational considerations, conceptualization of duties and towards future generations, lack of information, uncertainty and asymmetries of power” (Gabor, 2013, p. 301).

The standard economic approach deals with the intertemporal allocation of resources assuming that the utility function to maximize depends positively not only on the bundle of consumption of the present (person) generation, but in a certain degree also on the consumption—or on the utility—of the descendants (Arrow, 1973; Dasgupta, 1974; Solow, 1974). Then, a social discount rate is introduced to represent the relative degree of concern that the present generation bears for the next ones, so that the maximization of the present utility keeps balanced with the interests of future generations. In other words, the discount rate is supposed to prevent overconsumption behaviours of the present generation which could damage the future ones.

However, solving the intergenerational dilemma through a discount rate embedded in the utility function seems to be quite limiting for at least two reasons. With this approach, a lot depends on the adopted utility functional shape and on the specific value assigned to the social rate of discount. In particular, there is complete uncertainty about future generations' utility functions. Future people do not exist yet for definition and they cannot reveal their preferences (Beckerman, 2006; Parfit, 1984), therefore we cannot exactly know what is better for them (Barry, 1977). We can just conjecture it. Furthermore, there is no unanimity on how much exactly future generations should matter to our present decisions: in the environmental economics field, there is still a heated debate regarding the appropriate weight (social discount rate) we should assign to future generations' welfare (Nordhaus, 2008; Moore et al., 2004; Stern, 2008).

It is within this ambiguous utilitarian theoretical frame that we consider the social contract theory of John Rawls (Rawls, 1999) as a valid alternative to deal with distribution of resources between generations, and in particular with the climate change agreements. Indeed, with Rawls's social contract model, without discussing about social rates of discount or about future people's preferences, it seems to be possible to justify in a compelling way the idea that a closed set of self-regarding individuals (contemporaries) can take into consideration the interests of individuals left outside of the agreement (future generations).

3 John Rawls's social contract theory on allocation of resources between generations

John Rawls's theory is designed to identify the main principles which should shape societal institutions. In order to derive these principles, and assuming that individuals are instrumentally rational and mutually disinterested, that is they are not subject to any kind of social preference, Rawls proposes a procedure based on an impartial agreement. The impartiality in the choice of the principles is then guaranteed by a veil of ignorance, that is a tool which excludes the access to any particular information to the parties involved in the agreement (Rawls, 1999, pp. 118–123). In other words, while discussing about the principles.

“no one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like” (Rawls, 1999, p.118).

Basically, with the veil of ignorance, no one can take advantage of specific contingencies so as to design principles, norms or institutions which might favour her own particular position (Rawls, 1999, p. 17). On the contrary, everybody results equally represented since everybody has to choose in the same situation of perfect (mis)informational symmetry.

Although redistribution of resources between generations “subjects any ethical theory to severe if not impossible tests” (Rawls, 1999, p. 251), John Rawls (1999, 2001) does not miss to extend the mentioned contractarian model in order to include this topic. Thus, Rawls adds a further specification: “[t]he persons in the [agreement] have no information as to which generation they belong” (Rawls, 1999, p. 118, p. 254).

When moving from the intragenerational context to the analysis of principles for regulating allocation of resources between generations, Rawls also specifies that even though people are deprived of the information concerning the generation they belong, that is even if they ignore the historical and economic development of the society they represent, the parties behind the veil of ignorance are all contemporaries (Rawls, 1999, p. 121). Said otherwise, the people involved in the agreement belong to the same generational cohort, even if they are focused on identifying intergenerational principles.

In this way, Rawls substantially constrains his intergenerational decision-making model to the physically existing people (Dierksmeier, 2006). Indeed, according to Rawls, it would be unrealistic to conceive an agreement which gathers together all the possible generations of the human history: this all-inclusive approach would be too abstract³ (Attas, 2009, Rawls, 1999, Rawls, 2001a, 2001b; Tremmel, 2009). Notwithstanding the specification concerning the contemporaneity of the parties, Rawls’s social contract theory seems to be promising for dealing with the intergenerational allocation of resources (Tremmel, 2013): because of the veil of ignorance, the present generation loses its privileged (dictatorial) position towards future generations and can design “a path over time which treats all generations justly during the whole course of a society’s history” (Rawls, 1999, p. 257). In this way, it seems to be really possible to rationalize the idea of deriving intertemporal norms from a purely intragenerational context (Gabor, 2013; Heyd, 2009).

In Rawls’s opinion, the standard intragenerational distributive principle—the so-called difference principle—which requires to maximize the expectations of the worst-off, is not a suitable tool to deal with redistribution of resources between generations (Rawls, 1999, p. 56, p. 69, and p. 72). This principle, applied to the intergenerational context, could produce some undesirable consequences, like zero savings for future generations (Attas, 2009; Brandstedt, 2017; Buchanan, 1987; Dasgupta, 1974 and Rawls, 1999, p. 159 and pp. 253–255). Therefore, Rawls proposes the just saving principle as the normative rule to regulate distribution of resources across time: thus, “in following a just savings principle, each generation makes a contribution to those coming later and receives from its predecessors” (Rawls, 1999, p. 254, Gaspard & Gosseries, 2007, Gosseries, 2008, 2016; Paden, 1997). This, without the necessity to introduce any other-regarding preference towards descendants.⁴

Therefore, according to Rawls, every (present) generation is expected to give up a share of its own resources in order to pass it to the following generation. Nevertheless, Rawls does not describe in detail the peculiar features of the just saving principle—providing for example a specific saving rate. Instead, Rawls sketches some general ethical restriction

³ Rawls’s model has to be “understood as a purely hypothetical situation” (Rawls, 1999, p. 11), that is it coincides with the adoption of a particular perspective, and the agreement is conceived as a simple mental experiment (Marcon et al., 2020). For this reason, Finneron-Burns (2017) suggests that it is also possible to include in the veiled agreement possible future people, even though this way of proceeding opens other issues linked to the non-identity problem.

⁴ For exploring this issue see (Attas 2009; Brandstedt 2017; English 1977; Finneron-Burns 2017; Heyd 2009; Mathis 2009; Norton 1989; Rawls 2001a, 2001b; Wall 2003 and Wallack 2006).

which the contractual parties should take into account when defining the specific saving path (Rawls, 1999, pp. 255–6). However, these details are beyond the aims of the present reconstruction.

In synthesis, in Rawls's intergenerational theory, only contemporaries take part in the agreement and individuals have to choose an allocative configuration which does not affect exclusively their own welfare. Instead, each single generation is required to evaluate distributive principles that will produce effects also on third parties left outside of the agreement—future generations—who cannot take part in the contract, cannot make demands, cannot make objections, cannot threaten and cannot punish the actual decision-maker: this setting represents the structure of the modern climate change agreements, and “the ‘just savings’ framework seems to be the appropriate one for discussing the issues raised by climate change” (Heath, 2013, p. 35).

4 Previous experimental studies on Rawls's intergenerational theory

Many different theoretical approaches have been proposed to address the environmental issue in a Rawlsian perspective, going from considering health and environment as primary goods or including animals in the agreement setting (Gardiner, 2011b; Penn, 1990 and Thero, 1995) to running a third level agreement (Clements, 2015). In the same way, many experimental works aimed at testing empirically the assumptions or the conclusions of Rawls's social contract theory (Gaertner & Schokkaert, 2012). Nevertheless, to the best of our knowledge, there was only one single study that tested Rawls's intergenerational theory within an experimental setting (Wolf & Dron, 2020).

Wolf and Dron's (2020) design is very intuitive. In the baseline treatment, a common endowment is provided to a group of five players. Single players are then randomly assigned to a position within a sequential dictator game (Bahr & Requate, 2007 and Cason & Mui, 1998) representing the timeline, so that each player symbolizes a generation. Starting from the player occupying the first position, participants sequentially enter the dictator role and claim a share of the common endowment. The share that players withdraw during their dictator turn constitutes their final payment. The game continues until either the fourth player takes a decision (the fifth player is dummy player) or the common endowment is exhausted—in order to represent the “dictatorship of the present”, every player who enters the dictator role is allowed to claim the 100% of the (remaining) endowment, with the consequence that nothing is left to the following generations.

In one of the proposed treatments, Wolf and Dron (2020) introduce a preliminary stage where players have the chance to freely discuss and to agree about a rule to share the common endowment between the five generations. In particular, players are asked to do so behind a veil of ignorance of Rawlsian type, that is not knowing the position they will occupy in the actual sequential dictator game. After the players agree on a sharing rule, they are assigned to a position (generation) and they sequentially enter the dictator role exactly as in the baseline treatment described above.

However, Wolf and Dron's (2020) attempt has to be considered unsuccessful for two reasons. First of all, the veiled agreement did not produce a fairer distribution of the common endowment between the five generations compared to the baseline treatment where players could not communicate. Thus, despite individuals having the chance to discuss and to agree on a distributive rule, in the treatment with the veiled agreement players in privileged positions kept profiting of their favourable contingencies (Wolf & Wagner, 2016).

Second and even more important, the implemented experimental design is to be considered inconsistent with the intergenerational social contract theory such as formulated by John Rawls. Indeed, in Wolf and Dron's (2020) experiment, all the generations are put behind the same veil of ignorance, so as to stipulate an agreement between generations (Anderson, 2013 and Finneron-Burns, 2017). Instead, as it was explained above, Rawls is really clear and careful to specify that even though behind the veil of ignorance contractual parties assume an intergenerational perspective—since they do not know the generation of the history they belong to—all of them are contemporaries, that is all of them belong to the same generation.

Notwithstanding this inaccuracy in the design, Wolf and Dron's experiment provides a good basis to test Rawls's intergenerational theory, mainly because it follows a widespread practice in the experimental economic literature, that is to simulate generations assigning players to different positions in a sequential game. This practice has indeed occurred in trust games (Schotter et al., 2006), within public good games (Baggio et al., 2018 and Chaudhuri et al., 2006), with ultimatum games (Schotter et al., 2007) and with common pool resource games (Chermak et al., 2002 and Fisher et al., 2004).

5 Our experimental design and predictive hypothesis

We improved Wolf and Dron's (2020) setting and designed our game in order to take into account all the specification made by John Rawls about the intragenerational nature of the agreement for intergenerational principles: in our baseline treatment (BT), generations are not represented by single players but they are rather constituted by groups of three players.⁵

Here, it is immediately necessary to make a specification concerning the expression “generation” used in the context of our experiment. In particular, with the term “generation”, we mean a set of people who do not overlap with some of themselves in different moments of the timeline. Thus, while the intragenerational group setting distinguishes our experiment from the agreement between generations of Wolf and Dron (2020), the non-overlapping mechanism differentiates our model from contractarian theories like the one conceived by Joseph Heath (2013). Indeed, in Heath's model (2013, p. 35), “co-operation can be sustained intergenerationally through a system of indirect reciprocity”, so that in a repeated context, my good (and strategic) action at time $t, t+1, \dots, t+n-1$ can indirectly

⁵ In the context of the climate change agreements, the single participants of each group should be considered as representatives of states (Rawls 2001a, p. 10). However, this kind of interpretation creates an intersection between intergenerational and international justice and might trigger an entire debate on Rawlsian domestic distributive justice versus Rawlsian international distributive justice (Beitz 1999, Loriaux 2012, Nagel 2005; Pogge 1988, 2001, 2003 just to quote some). Indeed, according to John (Rawls 2001a), strong distributive justice holds within societies but not between them, so an agreement between states like the one Rawls (1999) conceives for the national closed systems seems not to be necessary, up to the point that Rawls (2001a) never mentions anything like an international (between states) agreement for intergenerational purposes. To the best of our knowledge Kim (2019) is the only one that tries to solve this Rawlsian impasse conceiving an international, intergenerational principle of justice. Nevertheless, with considering the players of our game as state representatives, our model might be accused of not reproducing accurately the Rawlsian (1999) intergenerational agreement model, because Rawls circumscribes the intergenerational problem exclusively to the single national closed systems. However, as it was shown in the Introduction, problems like the global warming cannot be solved by the virtuous action of single (even if just) states. Therefore, for pure policy purposes, we allow ourselves to rework Rawls's theory in the interpretation of the game model, being anyway aware of the mentioned imperfect overlap and the debate that this stance might generate. We are thankful to an anonymous reviewer for raising this interesting observation.

benefit myself at time $t+1, t+2, \dots, t+n$ through a cascade of good (and strategic) actions of an overlapping chain of players that follow the same type of scheme of indirect advantages.

However, this is not the case of our experiment. Since it has to represent the time bomb problem intrinsic in the climate change, in our model, the generations do not overlap: once a group of players (generation) takes a decision, it does not enter the game anymore. In other words, in our experiment, the same group of people does not play repeatedly in different moments of time, so there is no space for Heat's (strategic) indirect reciprocity. In this sense, our model represents the narrowest conceptualization of intergenerational cooperation because nobody can expect to be indirectly benefitted or reciprocated by her or his own good action taken in a previous moment of time.

After the groups of three subjects are formed, each group is randomly assigned to a position on chains (sequences) of different lengths. The shortest chain is made of one group (generation), the longest one of a sequence of five groups (generations). This is another expedient we introduced as compared to the Wolf and Dron's (2020) design: the indeterminateness of the chain length. Despite not being a unanimous practice (Wolf & Dron, 2020 and Kamijo et al., 2017), this uncertainty element is anyway a common feature—considered also more realistic—in experiments of this kind (Fischer et al., 2004; Hauser et al., 2014). Starting then from the first generation, players are asked to play a group sequential dictator game.

The decision-making task for the dictator group is then designed as follows. The first group (generation) of each chain starts with a common endowment of €21. Each of the three participants of the group has to decide individually how much money to withdraw from the common pool, choosing an integer value between €0 and €7.⁶ The amount that each player claims for himself in this stage constitutes his individual final payment. After a player makes a choice and before he is revealed, the outcome of the group he is asked to guess⁷ the decisions taken by the other two players of his group.

After all the three players make their choice, one of the two following scenarios occurs:

- if the common pool is left with at least €6 in total, the chain continues and the next group in the sequence enters the decision-making phase, becoming de facto the new present generation. The common endowment is refilled up to the initial value of €21 and the active generation faces the same identical decision-making problem described so far;
- if in the common pool the players leave €5 or less, the resource is not renewed and the chain breaks up, with the consequences that all the following generations cannot take any decision and do not get paid (except for the show up fee of €3).

⁶ Given the structure of the game players belonging to the same group are in a symmetric position. Therefore, within this design, one of the two problems linked to the unsuccessful international agreements on climate actions, that is the distribution of costs and the self-serving bias (Kriss et al., 2011), is put aside. However, this simplification does not make the experiment less relevant in disentangling the international agreements on the reduction of greenhouse gases: if nations cannot reach an agreement in a situation of perfect symmetry, *a fortiori* we cannot expect a widely shared agreement in the case of asymmetric costs and benefits.

⁷ The guess has an incentivized structure: the players with the best guess in each group were rewarded with €2 extra. Given the symmetric roles, we adopted a simple sum of absolute distances between the guess and the actual choices to determine the player(s) with the best guess (the one(s) with the lowest sum of distances).

Basically, our experiment is a public good game with a threshold: only reaching this last one the effect is produced. However, compared to previous experiments which used this setting to enquire climate change issues (Brick and Visser, 2014, Milinski et al., 2008 and Tavoni et al., 2011), in our game, the loss is not imposed to the same participants who take the decisions, but to another pool of subjects.

An experimental session lasts up to the point that all chains either get to their natural end or interrupt. The minimum threshold⁸ mechanism simulates the extreme threat embedded in the climate change: if the present generation overexploits the environment, it does so at the expense of future generations. On the contrary, if individuals in the present generation coordinate for leaving a minimum amount of (renewable) resources for the next generation, the last one can enjoy the same opportunities as the former (Brundtland Report, 1987).

The veil treatment (VT) adds to the baseline game a preliminary stage where the three players of every single group can reach an agreement before being placed in a position of a chain. Thus, consistently with Rawls's theoretical framework, while the groups of contemporaries vote for a principle aimed at managing the appropriation of the common resources, they do so not knowing the generation (position) they belong in the history (in the chain). In particular, at the beginning of the VT, every group is asked to unanimously agree on one of the two following principles:

- *Continuation of the chain*: Each participant of my group should withdraw a maximum of €5 from the common account, ensuring in this way a minimum total saving of €6 that allows the chain to continue.
- *Interruption of the chain*: Having the possibility to do it, each participant of my group should withdraw from the common account more than €5, even if that means interrupting the chain and leaving the other generations with €0.

The groups that reach an agreement are positioned on a chain. Instead, the groups which do not reach an agreement on the rule by the sixth round do not take part in the sequential dictator game. The players of these discarded groups are paid only the show up fee of €3.

Given the theoretical frame and the experimental design described so far, it is possible to formulate the predictive hypothesis of our intergenerational game. Our first hypothesis concerns the baseline treatment (BT) and it follows from the standard economic assumptions: the sub-perfect equilibrium is represented by the triple (€7, €7, €7) for every generation in any chain, therefore, players in the first generation are expected to appropriate the total available endowment, leaving no resources in the common pool. This behaviour is expected to undermine the chains continuation.

H1 In the BT, the players of the first generation will claim €7 each, breaking up all the chains.

Our second hypothesis concerns the veil treatment (VT) and directly follows from the Rawlsian theory. More precisely, behind the veil of ignorance, players should agree on a norm representing the just saving principle, according to which “each generation

⁸ Except for special cases (Bardsley 2008; Cherry et al., 2002; List 2007), the threshold established in our experiment (6/21 in total or 2/7 individually) mirrors the average amount of money left by active players (28.5%) to dummy players in dictator games (Engel, 2011).

makes a contribution to those coming later and receives from its predecessors” (Rawls, 1999, p. 254). Therefore, in the VT, groups are expected to agree on the rule which constrains a pure self-interest behaviour.

H2 In the voting phase of the VT, the groups will agree on the rule representing the Rawlsian just saving principle, that is “continuation of the chain”.

The next hypothesis on the VT deserves a brief theoretical digression, also related to the climate change agreements. In general, the agreement behind the veil of ignorance is not conceived by Rawls as binding. Thus, although the principles are the result of a fair procedure, once the veil is lifted and the groups are assigned to a chain position, a dictatorship of the present generation over the future ones keeps being a concrete possibility (Wolf & Dron, 2020 and Wolf & Wagner, 2016). In other words, the voting outcome is not automatically implemented, like it was done in other intergenerational experiments (Hauser et al., 2014; Kamijo et al., 2017).

Indeed, the practice of the mechanical enforcement of the agreement is considered unrealistic. First of all, the actual individual behaviour can deviate from the abstract voting expression. Second, even though the authors (Hauser et al., 2014) justify the binding vote as a good proxy for informal institutions which usually enforce co-operative attitudes (like punishments or rewards), those enforcement mechanisms work only when there are repeated interactions among the same subjects, so for example paying a cost now (punishing) can generate long-term benefits. But this is not the case of their experiments. Instead, the possibility for voluntary compliance after the vote represents a realistic setting, analogous to one of the international agreements on emissions reduction, since they are not enforced (nor enforceable) by any authority, either formal or informal.

In the experimental field, the compliance to non-binding distributive principles was studied through a double dictator game, the so-called exclusion game. The exclusion game is a resource allocation game which includes a preliminary voting stage behind a veil of ignorance that prevents the parties from knowing their role (dictator or dummy) in the actual game (Faillo et al., 2014; Rawls, 1963; Rawls, 1999; Sacconi & Faillo, 2010; Sacconi et al., 2011; Tammi, 2011). In the game, the compliance to the agreement, conditioned on the belief of other players’ compliance, was high even in those cases in which players agreed on a counter-maximizing rule (Faillo et al., 2014, p. 242). This attitude has been modelled through the conformist social preferences model (Grimalda & Sacconi, 2005) and based on the Rawlsian concept of the sense of justice (Rawls, 1999).

Thus, we can expand the same idea of the sense of justice to the intergenerational context. Indeed, in our design, the active generation is not constrained by any enforcement mechanism to apply the outcome of the agreement reached behind the veil of ignorance, because the compliance is left to an individual choice. This is a realistic structure since in the real world we have no concrete institutions which can constrain the present generation to care about the future ones, even if the intergenerational norms are the outcome of a fair procedure. Therefore, we expect compliance to be high.

H3 In the VT, individuals will comply with the intragenerational agreement, believing that the other components of the group will comply with the norm.

The last hypothesis becomes a logical sum of the previous two: if in the VT, the groups agree on an intergenerational sustainable behaviour—continuation of the chain—and if the players comply with the chosen norm, the chains will continue up to their last generation.

H4 Compared to the BT, in the VT, a higher number of chains will continue until their natural end.

The just mentioned hypothesis is also the result of two other more specific but substantially equivalent sub-hypothesis.

H4a The proportion of people claiming a €5 or less will be higher in the VT than in the BT.

H4b The average individual claim will be lower in the VT than in the BT.

6 Data description

All the experimental sessions took place in the spring of 2019 in the Computable and Experimental Economics Laboratory (CEEL) of the University of Trento. They were run using the free software for economic experiments z-Tree (Fischbacher, 2007). All participants who took part in the experiment showed up voluntarily after a public call. The experiment involved a total of 141 participants (81 in the BT and 60 in the VT). On average, the participants were 22, 54% of them were females and 46% of the total subjects were enrolled in programmes related to the economic discipline. Participants were privately paid in cash at the end of each session, and on average, they earned about €6 (show up fee of €3 included). Each experimental session lasted at a maximum of 45 min.

In the experimental laboratory, the participants were randomly assigned to a computer terminal. All the emplacements were isolated by separation walls to avoid communication. The participants were given paper instructions. The instructions were also read aloud to ensure common knowledge. In the final questionnaire, participants declared that the provided instructions were very clear (4.6 on average in a range varying between 1 = not clear at all and 5 = very clear). After reading the instructions and before starting the experiment, in the BT (VT), 4 (6) control questions about the structure of the game were asked.

In the BT, we run four sessions for a total of eight chains, and in the VT, we run three sessions with two chains each (Fig. 1). In the BT, six chains out of eight (75%) broke up after the choices made by the players belonging to the first generation. In the two remaining chains (session 1 and 3), despite having only two generations, the second generations essentially interrupted the sustainable behaviour initially adopted by the first generations. Thus, in the BT, out of 81 participants, only 30 (37%) of them played an active role in the game taking an actual decision. These data are consistent with the predictions of **H1**, because in the BT, all the chains broke up at the very beginning.

From the common endowment of €21, in the BT, active players withdraw on average €5.30 (s.d. 1.09) believing that the other two players in the group would claim on average €5.20 (s.d. 1.22). These last data (represented in Fig. 2) show an interesting empirical regularity which was not taken into account by the predictive hypothesis: generations in the BT waste resources, because the active groups left in the common pool at least €2. In general, the active players left in the pool about 1.70€ on average, for a total

BASELINE TREATMENT
(no vote)

GENERATION SESSION	GEN 1	GEN 2	GEN 3	GEN 4	GEN 5
Session 1	(4,4,4)	(4,5,7)			
	(5,7,7)	X	X	X	X
Session 2	(4,5,7)	X			
	(4,5,7)	X	X	X	
Session 3	(5,5,5)	(5,5,6)			
	(4,5,7)	X	X	X	X
Session 4	(5,5,6)	X	X	X	
	(5,5,7)	X	X		

VEIL TREATMENT
(the vote outcome is reported above the decision)

GENERATION SESSION	GEN 1	GEN 2	GEN 3	GEN 4	GEN 5
Session 1	continuation (5,5,5)	continuation (5,5,7)			
	continuation (4,4,5)	continuation (5,5,7)	continuation X	continuation X	continuation X
Session 2	continuation (5,5,7)	continuation X			
	continuation (4,5,7)	interruption X	continuation X	continuation] X	
Session 3	continuation (5,5,7)	continuation X	continuation X	continuation X	
	interruption (7,7,7)	continuation X	continuation X		

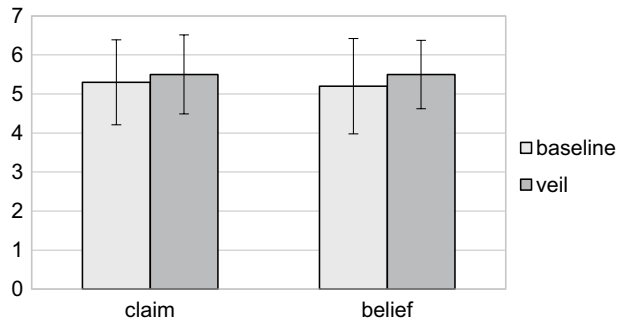
Legend

- = sustainable decision, chain continuation
- = overconsumption, chain broken up
- X = no decision taken

Fig. 1 Representation of the chains and of the decisions of the groups for each session

of €50 that remained part of the experimenter budget for running other experimental sessions—indeed, in the game, this amount of money is basically destroyed since it does not allow the chains to continue anyway.

Fig. 2 Individual average claim and average belief (€) per treatment



Seen from another point of view, we can claim that some participants—about 70% of the active players in the BT—revealed altruistic preferences. Indeed, these players individually renounced to consume a share of their maximum earning without the possibility for being directly reciprocated by the players belonging to the next generations. However, their good purposes were nullified by the actions of a minority of players who did not do their part in the fair management of the common resources.

As far as the preliminary voting stage of the VT is concerned, all groups reached a unanimous agreement on the rule by the second round. This proves that the veil of ignorance is a powerful tool to coordinate players on climate change issues (Kriss et al., 2011). More specifically, 18 out of 20 groups unanimously agreed on the sustainable rule labelled “continuation of the chain” (Fig. 1). Therefore, **H2** is supported by the data too, because most of the participants—54 players, representing the 90% of the subjects involved in the VT—behind the veil of ignorance voted for the just saving principle.

In the VT, only 24 out of 60 participants (40%) played in the active role of the sequential dictator game taking an actual decision. However, after the agreement phase, 19 out of 24 (80%) participants who entered the decision-making phase complied with the rule chosen behind the (laboratory) veil. Therefore, the experimental data support also **H3**, because the majority of participants followed the approved rule even in those many cases where that was against their own material interest.

Nevertheless, although the majority of participants chose the just saving principle and most of them complied with it, in the VT, no one chain continued beyond the second generation, showing no substantial differences with the BT (Fig. 1). Therefore, according to the empirical data, **H4** has to be rejected because in the VT no one chain got to its natural end as predicted.

That **H4** is not supported by data follows also from the rejection of its sub-hypothesis. **H4a** is rejected since between the two treatments, the proportion of the active players claiming an amount of €5 or less is exactly the same and corresponding to 0.7. In the same way, **H4b** is rejected because in the VT, players withdraw on average €5.50 (s.d. 1.01), an amount actually higher than in BT (Fig. 2). However, this behaviour left in the experimenter budget a total of €21, since in the VT, active players decided to leave in the common pool €1.50 on average. Also, the average belief on others’ claims is higher and corresponding to 5.50€ (s.d. 0.88), showing no significant difference with the BT.

The undifferentiated behaviour between the two treatments, which leads to the rejection of **H4b** and then of **H4**, can be deduced also through a simple linear model (1) where we regress the individual claim on the *treatment* (dummy variable, where 0=BT), on the *generation* (position in the chain) and on the average *belief* (in €) – plus a series of control demographical factors (Table 1). We clustered the standard errors per single groups.

Table 1 Determinants of individual claims, OLS regression

Variables ^a	Treatment OLS	Pooled OLS
Treatment	0.18 (0.296)	
Generation	0.53 (0.267)*	0.53 (0.266)*
Belief	0.48 (0.113)***	0.50 (0.119)***
Instructions	0.35 (0.134)***	0.38 (0.125)***
Experiments	0.04 (0.005)**	0.03 (0.007)***
Age	-0.07 (0.080)	-0.08 (0.066)
Gender	-0.39 (0.264)	-0.37 (0.268)
Nationality	0.75 (0.307)**	0.81 (0.265)***
Field of studies	-0.40 (0.218)	-0.07 (0.206)
Years of studies	0.11 (0.108)	0.13 (0.106)
Risk	0.04 (0.037)	0.04 (0.038)
Family income	0.24 (0.110)**	0.25 (0.111)**
Economic situation	0.50 (0.213)**	0.45 (0.200)**
Adjusted R-squared	0.6013	0.6052

Standard error in parentheses, n. of observations 51^b, * significant at 90%, ** at 95%, *** at 99%

^a The variable *instructions* refers to the comprehensibility of the initial instructions (ranging from 0 to 5), *experiments* is the number of previous experiments, *age* the personal age, *gender* is a binary dummy for the gender (male=0), *nationality* is a dummy variable for the nationality (0=Italian, 1=other), *field of studies* and *year studies* refer to the undertaken studies (the first one is a dummy variable, 1=economic disciplines, 0=other fields, the second one is a discrete variable), *risk* is a subjective statement about the risk attitude (1 risk adverse, 10 risk seeker), *income* and *economic situation* measure, respectively, the yearly family income and the current economic status (both are discrete variables ranging from 1, very low/bad, to 5, very high/excellent)

^b The number of observations is 51 instead of 54 because we could not collect the demographic statements of three participants in the BT, so the models are calculated without those data

$$claim_i = \alpha + \beta_1 treatment_i + \beta_2 generation_i + \beta_3 averagebelief_i + controls_i + \varepsilon_i \quad (1)$$

From the econometric estimation of the model, we can see that there is no statistically significant difference in the average claim between the two treatments. Among the main predictors, only the average belief on the others' withdrawal and the position in the chain had a strong and significant effect on the individual choice: in particular, a higher expectation of €1 on the others' average withdrawal increases the average individual claim of about €0.5; belonging to the second generation (the last generation who actively played) also increases the average claim of €0.5.

7 Comment to the data

Our experiment basically shows that a laboratory veil of ignorance facilitates the achievement of a unanimous agreement for the fair management of resources across generations. At the same time, the individual compliance to this scheme, despite being high, was not sufficient to allow the chains to survive significantly longer than in the baseline treatment, where there was no agreement phase.

This result highlights what is considered a limit of our experimental design: even though the compliance rate to the just saving principle was very high—over 75%—and consistent with the previous empirical evidence provided by the exclusion game, this percentage seems to be still too low compared to the 100% of compliance required by the structure of the game to allow a chain to continue. In other words, in our experiment, a society (a chain) can move to the next generation only if the compliance to the continuation principle is absolute. Seen from another perspective, one single withdrawal over the threshold (€6 or €7) is sufficient to undermine the entire scheme of intragenerational co-operation, and this is equivalent to giving to every participant a veto power on the common sustainable action.

This consideration helps to understand better why our institutional mechanism of the veil of ignorance did not work properly as predicted. Focusing on the distribution of claims in the two treatments (Fig. 1), we can notice how in 12 out of 14 groups which interrupted the chain, the continuation of the chain fails because one single player out of three claims more than €5. This veto power has a strong effect even though about 13% (5 out of 39) of the active players tried to anticipate this behaviour and to compensate for this risk claiming less than €5. In synthesis, expecting perfect voluntary compliance (or a huge sacrifice from single individuals to compensate the excesses of others) seems to be utopian, as much in an experiment as in the real world.

Left aside the veto issue, another reason why we do not observe differences between the two treatments is because that in the BT, the active players are already too altruistic. This makes difficult to observe any significative improvement in the VT. Thus, comparing the choice triples of the two treatments (Fig. 1), we notice that the agreement phase of the VT is probably too faint to produce any effect on the behaviour of those “hard” participants who withdraw a high amount from the common pool whatever is the underlying institutional mechanism.

This last criticism opens then a broader concern about the functioning of the Rawlsian veil of ignorance and the sense of justice for intergenerational principles. Although according to the data, the belief-based behaviour (Table 1) seems to strengthen the literature stream on the Rawlsian sense of justice in the laboratory (Faillo et al., 2014), it should be inquired better whether the sense of justice at the basis of the compliance to an impartial agreement can be triggered when a set of people *X* agrees to undertake an action towards a set of people *Y* and different from *X*. As clarified in Sect. 2, Rawls’s theory establishes that a generation of contemporaries evaluates distributive principles that will produce effects on third parties—future generations—not formally present in the agreement.

Thus, the (between groups) social contract on intergenerational principles is very different from the standard (within groups) distributive issue where a set of individuals has to agree on how to share resources between themselves: whether I have to agree with you on how to split €10 between ourselves or whether I have to agree with you on how to split €10 between ourselves and a third person who has no voice on the issue are two extremely different decision-making situations. In the latter case, we are supposed to take into consideration some parties who remain excluded from the contract but who are direct stakeholders

of our decision. Furthermore, in the intergenerational context, after the veil is dropped, the parties involved in the agreement and the external stakeholders do not enter a mutually beneficial co-operative scheme as it happens for the standard intragenerational case (Rawls, 1999, p. 4). These differences have to be born in mind when extending the validity of the conformist preference model and its underlying idea of the sense of justice to the joint commitment for distribution of resources between generations (Gilbert, 2013).

8 Conclusions

Climate change agreements and compliance to their prescriptions are one of the most challenging issues of the last decades. If nations do not seriously start co-operating to stop the growth of greenhouse gas emissions in the atmosphere, humankind will soon experience some tragic consequences (Hardin, 1968). With our laboratory experiment based on John Rawls's social contract theory, we addressed the systemic lack of compliance to the international climate change agreements. We focused on voluntary compliance because currently there are no transnational monitoring and sanctioning institutions, so enforced voting procedures (Hauser et al., 2014; Kamijo et al., 2017) are not considered representative of the real context.

We have seen that two are the main hurdles that prevent countries to produce effective agreements concerning the reduction of greenhouse gas emissions (Gardiner, 2011a). We left aside the problem regarding the distribution of costs between nations, assuming symmetrical situations between the participants belonging to the same generation. However, this simplification does not make our experiment less useful in order to solve the compliance issue: if nations are not able to share and to follow a principle in a situation of perfect symmetry, a fortiori we cannot expect a widely shared agreement in the case of asymmetric costs and benefits. Therefore, our experiment constitutes an important step in understanding better the structural problems of the international agreements on climate actions.

We can summarize the contribution of our study through two main conclusions. First, the intragenerational veil of ignorance for intergenerational purposes proves to be very powerful at inducing the groups to formally agree on a sustainable rule for the resources division. This result should not be considered trivial if we remind that we can count about 15 different international environmental agreements (and an indefinite number of bargaining rounds) in less than 30 years, as a proof that unanimously agreeing on a resolute principle for acting against the global warming is not easy at all. Thus, a mechanism of impartiality seems to play an important role for limiting the self-serving bias dominating this type of agreements. A next step might be to explore this tool in a situation where the distribution of costs and benefits is not symmetrical.

Nonetheless, and this is the second important result, the veil of ignorance did not produce any significative effect in terms of voluntary compliance, and therefore on the continuations of the chains, compared to the baseline treatment. In particular, even a small monetary incentive to deviate from the agreement—€2 more in the own pocket—was sufficient to lead to a chain break, implying a potential loss of € 75 for all the following groups representing future generations. This result shows that the Rawlsian theory, as represented by our setting, does not really help us to solve the impasse of compliance to the international climate actions agreements.

However, this result has also to take into account the structural limits of our game which basically required a perfect compliance to the agreement in order to allow the generational

chains to continue. Moreover, given that the generations do not overlap, we have to keep in mind the dissimilarities between the exclusion game and our agreement on intergenerational principles: the agreement of the former includes all the stakeholders while the latter leaves outside the least advantaged subjects, that is the future generations. This might interfere with the mutual expectations of compliance and therefore with the sense of justice, exactly because after the veil is dropped, the parties do not enter a mutually advantageous scheme of co-operation, nor a dynamic of sequential indirect reciprocity (Heath, 2013). Indeed, our model is based on the narrowest conceptualization of intertemporal co-operation, so that a generation cannot get any direct or indirect advantage from benefitting the following one. Also, this very austere modelling might have contributed to the less favourable results compared to more optimistic models (Heath, 2013).

Thus, as it happens with the real international agreements on climate actions, even if everybody formally agrees that we have to care about future generations, at individual level, other small incentives prevail, even in front of a huge cost. The big issue on how to structure self-enforcing international agreements on climate actions keeps open, but with some insights for the future attempts. For example, a more moderate compliance mechanism—but which avoids public good dynamics—or the possibility of introducing an imaginary person from the future which can activate the memory of the agreement in the choice phase (Kamijo et al., 2017; Hara et al., 2019) could help to structure an effective agreement. These mechanisms could foster compliance and prevent future generations from bearing the costs of our overconsumption behaviour today.

Otherwise, another possibility is to frame the game such that the link with the climate change issue is more evident. In this case, the interruption of the generational chain should not be a direct consequence of the overconsumption behaviour, but of the by-product of the overconsumption, that is the emissions. This means introducing a function that in the game represents the externality (Ostrom 1994) of the consumption choices. The model might be improved also introducing an overlapping mechanism between generations (Heath, 2013), so that long-run goals that we tried to represent with our experiments can be achieved through an indefinite number of short-run co-operative steps.

Another alternative that future research might explore is the introduction in the agreement discussion of a climate tax proportioned to the generated emissions, as suggested by Puauschunder (2020). However, if we have to imagine such a tax as binding, this is equivalent to introduce some international monitoring institutions with some specific sanctioning powers towards the countries that do not meet the agreed targets of reduction or do not comply with the due tax payments. In this case, we have to be aware that this hypothesis might complicate even more the agreement phase, because parallelly with choosing the institution protecting the future generations, states should agree on a controlling mechanism for such an institution.

Acknowledgements We are particularly thankful to Lucas Stanczyk for inspiring us to understand the problems of the international climate agreements from a Rawlsian perspective. The work benefited also from helpful comments provided by Stefano Castriota, Giacomo Degli Antoni, Pedro Francés Gómez, Roberto Tamborini and Tatsuyoshi Saijo and by other three anonymous reviewers.

Funding Open access funding provided by Università degli Studi di Trento 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

- Anderson, M. W. (2013). Intergenerational bargains: negotiating our debts to the past and our obligations to the future. *Futures*, *54*, 43–52.
- Arrow, K. J. (1973). Rawls's principle of just saving. *The Swedish journal of economics*, *75*(4), 323–335.
- Attas, D. (2009). A transgenerational difference principle, intergenerational justice, 189–218. In Axel Gosseries (Ed.), *and Lukas H.* Oxford: Meyer, Oxford University Press.
- Baggio, M., & Mittone, L. (2019). Grandparents matter: perspectives on intergenerational altruism and a pilot intergenerational public good experiment. *Homo Oeconomicus*, *35*(4), 1–22.
- Bahr, G., & Requate, T. (2007). *Intergenerational fairness in a sequential dictator game with social interaction*. Germany: Kiel University.
- Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*, *11*(2), 122–133.
- Barry, B. (1977). Justice between generations. In: Hacker P, Raz J (eds) *Law, morality, and society. Essays in honour of H.L.A. Hart*. Clarendon Press, Oxford.
- Beckerman, W. (2006). The impossibility of a theory of intergenerational justice, *Handbook of intergenerational justice*, 53–71, edited by Tremmel J. C., Edward Elgar.
- Beitz, C. R. (1999). *Political theory and international relations, with afterward*. Princeton University Press.
- Brandstedt, E. (2017). The savings problem in the original position: assessing and revising a model. *Canadian Journal of Philosophy*, *47*(2–3), 269–289.
- Brekke, K. A., & Johansson-Stenman, O. (2008). The behavioural economics of climate change. *Oxford review of economic policy*, *24*(2), 280–297.
- Brick, K., & Visser, M. (2015). What is fair? An experimental guide to climate negotiations. *European Economic Review*, *74*, 79–95.
- Report, B. (1987). *Our common future*. Report of the world commission on environment and development.
- Buchanan, J. M. (1987). The constitution of economic policy. *The American economic review*, *77*(3), 243–250.
- Burke, E. (1993). *Reflections on the revolution in France*, edited with an introduction of Mitchell L. G., Oxford University Press. Oxford
- Cason, T. N., & Mui, V. L. (1998). Social influence in the sequential dictator game. *Journal of mathematical psychology*, *42*(2–3), 248–265.
- Chaudhuri, A., Graziano, S., & Maitra, P. (2006). Social learning and norms in a public goods experiment with inter-generational advice. *The Review of Economic Studies*, *73*(2), 357–380.
- Chermak, J. M., & Krause, K. (2002). Individual response, information, and intergenerational common pool problems. *Journal of Environmental Economics and Management*, *43*(1), 47–70.
- Cherry, T. L., Frykblom, P., & Shogren, J. F. (2002). Hardnose the dictator. *The American Economic Review*, *92*(4), 1218–1221.
- Clements, P. (2015). Rawlsian ethics of climate change. *Critical Criminology*, *23*, 461–471.
- Dasgupta, P. (1974). On some problems arising from professor rawls' conception of distributive justice. *Theory and Decision*, *4*(3), 325–344.
- Dierksmeier C. (2006). John Rawls on the rights of future generations, *Handbook of intergenerational justice*, 72–85, edited by Tremmel J. C., Edward Elgar.
- Dinda, S. (2004). Environmental kuznets curve hypothesis: a survey. *Ecological economics*, *49*(4), 431–455.
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, *14*(4), 583–610.
- English, J. (1977). Justice between generations. *Philosophical Studies*, *31*(2), 91–104.
- Faillo, M., Ottone, S., & Sacconi, L. (2014). The social contract in the laboratory: an experimental analysis of self-enforcing impartial agreements. *Public Choice*, *163*(3–4), 225–246.
- Finneron-Burns, E. (2017). The intergenerational original position. *Social Theory and Practice*, *43*(4), 805–823.
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics*, *10*(2), 171–178.

- Fischer, M. E., Irlenbusch, B., & Sadrieh, A. (2004). An intergenerational common pool resource experiment. *Journal of environmental economics and management*, 48(2), 811–836.
- Gabor, A. (2013). The savings principle and inter-generational justice. *European Journal of Science and Theology*, 9(2), 299–308.
- Gaertner, W., & Schokkaert, E. (2012). *Empirical social choice: questionnaire-experimental studies on distributive justice*. Cambridge University Press.
- Gardiner, S. M. (2009). A Contract on Future Generations?, Intergenerational justice, 77–118. In Axel Gosseries (Ed.), and Lukas H. Oxford: Meyer, Oxford University Press.
- Gardiner, S. M. (2011a). *A perfect moral storm: the ethical tragedy of climate change*. Oxford University Press.
- Gardiner, S. M. (2011b). Rawls and climate change: does Rawlsian political philosophy pass the global test? *Critical Review of International Social and Political Philosophy*, 14(2), 125–151.
- Gardiner, S. M., & Weisbach, D. A. (2016). *Debating climate ethics*. Oxford University Press.
- Gaspard, F., & Gosseries, A. (2007). Are generational savings unjust? *Politics, Philosophy & Economics*, 6(2), 193–217.
- Gilbert, M. (2013). *Joint commitment: How we make the social world*. Oxford: Oxford University Press.
- Gosseries A. (2008). Teorie della giustizia intergenerazionale: Una sinopsi, Notizie di Politeia. - ISSN 1128–2401. - 24:91, pp. 7–26.
- Gosseries, A. (2016). La Cuestión Generacional y la Herencia Rawlsiana. *Revista Electrónica Instituto de Investigaciones Jurídicas y Sociales AL Gioja*, 8, 71–90.
- Grimalda, G., & Sacconi, L. (2005). The constitution of the not-for-profit organisation: reciprocal conformity to morality. *Constitutional Political Economy*, 16(3), 249–276.
- Hara, K., Yoshioka, R., Kuroda, M., Kurimoto, S., & Saijo, T. (2019). Reconciling intergenerational conflicts with imaginary future generations: Evidence from a participatory deliberation practice in a municipality in Japan. *Sustainability Science*, 14(6), 1605–1619.
- Hardin, G. (1968). The tragedy of the commons. *Science*, 162, 1243–1248.
- Hauser, O. P., Rand, D. G., Peysakhovich, A., & Nowak, M. A. (2014). Cooperating with the future. *Nature*, 511(7508), 220.
- Heath, J. (2013). The structure of intergenerational cooperation. *Philosophy & Public Affairs*, 41(1), 31–66.
- Heyd, D. (2009). A Value or an obligation? Rawls on justice to future generations, 167–188. In Axel Gosseries (Ed.), and Lukas H. Oxford: Meyer, Oxford University Press.
- Isaac, R. M., & Walker, J. M. (1988). Group size effects in public goods provision: The voluntary contributions mechanism. *The Quarterly Journal of Economics*, 103(1), 179–199.
- Kamijo, Y., Komiya, A., Mifune, N., & Saijo, T. (2017). Negotiating with the future: Incorporating imaginary future generations into negotiations. *Sustainability science*, 12(3), 409–420.
- Kim, H. (2019). An extension of Rawls's theory of justice for climate change. *International Theory*, 11(2), 160–181.
- Klinsky, S., & Dowlatabadi, H. (2009). Conceptualizations of justice in climate policy. *Climate Policy*, 9(1), 88–108.
- Kriss, P. H., Loewenstein, G., Wang, X., & Weber, R. (2011). Behind the veil of ignorance: Self-serving bias in climate change negotiations. *Judgment and Decision Making*, 6(7), 602–615.
- Kuznets, S. (1955). Economic growth and income inequality. *The American economic review*, 45(1), 1–28.
- Lamont, J., and Favor, C., "Distributive Justice", online resource, The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.), entered August 2020.
- Lange, A., Löschel, A., Vogt, C., & Ziegler, A. (2010). On the self-interested use of equity in international climate negotiations. *European Economic Review*, 54(3), 359–375.
- List, J. A. (2007). On the interpretation of giving in dictator games. *Journal of Political economy*, 115(3), 482–493.
- Loriaux, S. (2012). Fairness in international economic cooperation: moving beyond Rawls's duty of assistance. *Critical Review of International Social and Political Philosophy*, 15(1), 19–39.
- Marcon, L., Frances-Gómez, P., & Faillo, M. (2020). Does impartial reasoning matter in economic decisions? An experimental result about distributive (un) fairness in a production context. *Theoria: An International Journal for Theory, History and Foundations of Science*, 35(2), 217–233.
- Mathis, K. (2009). Future generations in John Rawls' theory of justice. *Archiv für Rechts-und Sozialphilosophie*, 95(1), 49–61.
- Meyer, L., "Intergenerational Justice", online resource, The Stanford Encyclopedia of Philosophy, Edward N. Zalta (ed.), entered August 2020.
- 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.

- Moore, M. A., Boardman, A. E., Vining, A. R., Weimer, D. L., & Greenberg, D. H. (2004). "Just give me a number!" Practical values for the social discount rate. *Journal of Policy Analysis and Management*, 23(4), 789–812.
- Nagel, T. (2005). The problem of global justice. *Philosophy & public affairs*, 33(2), 113–147.
- Nordhaus, W. D. (1993). Reflections on the economics of climate change. *Journal of Economic Perspectives*, 7(4), 11–25.
- Nordhaus, W. D. (2008). *A question of balance: Weighing the options on global warming policies*. London: Yale University Press.
- Norton, B. G. (1989). Intergenerational equity and environmental decisions: A model using Rawls' veil of ignorance. *Ecological Economics*, 1(2), 137–159.
- Ostrom, E., Gardner, R., Walker, J., & Walker, J. (1994). *Rules, games, and common-pool resources*. University of Michigan Press.
- Puaschunder, J. (2020). Mapping climate justice. In *Governance & Climate Justice* (pp. 23–38). Palgrave Macmillan, Cham.
- Paden, R. (1997). Rawls's just savings principle and the sense of justice. *Social Theory and Practice*, 23(1), 27–51.
- Parfit, D. (1984). *Reasons and Persons*. Oxford University Press.
- Penn, J. (1990). Towards an ecologically-based society: a Rawlsian perspective. *Ecological Economics*, 2(3), 225–242.
- Pogge, T. W. (1988). Rawls and global justice. *Canadian Journal of Philosophy*, 18(2), 227–256.
- Pogge, T. W. (2001). Rawls on international justice. *The Philosophical Quarterly*, 51(203), 246–253.
- Pogge, T. W. (2003). The incoherence between Rawls's theories of justice. *Fordham L. Rev.*, 72, 1739.
- Rawls, J. (1963). The sense of justice. *The Philosophical Review*, 72(3), 281–305.
- Rawls, J. (1999). *A theory of justice* (Revised). Cambridge: Harvard University Press.
- Rawls, J. (2001a). *Justice as fairness: A restatement*. Harvard University Press.
- Rawls, J. (2001b). *The Law of Peoples: with, the idea of public reason revisited*. Harvard University Press.
- Sacconi, L., & Faillo, M. (2010). Conformity, reciprocity and the sense of justice: how social contract-based preferences and beliefs explain norm compliance: the experimental evidence. *Constitutional Political Economy*, 21(2), 171–201.
- Sacconi, L., Faillo, M., & Ottone, S. (2011). Contractarian compliance and the sense of justice': a behavioral conformity model and its experimental support. *Analyse & Kritik*, 33(1), 273–310.
- Sayegh, A. G. (2020). Moral duties, compliance and polycentric climate governance. *International Environmental Agreements: Politics, Law and Economics*, 20(3), 483–506.
- Schotter, A., & Sopher, B. (2006). Trust and trustworthiness in games: An experimental study of intergenerational advice. *Experimental Economics*, 9(2), 123–145.
- Schotter, A., & Sopher, B. (2007). Advice and behavior in intergenerational ultimatum games: An experimental approach. *Games and Economic Behavior*, 58(2), 365–393.
- Smith, A. (1976). *The theory of moral sentiments*. Edited D. D. Raphael and A. L. Macfie. Clarendon Press.
- Smith, A. (1994). *An inquiry into the nature and causes of the wealth of nations*. Edited by Edwin Cannan. The Modern Library. New York.
- Solow, R. M. (1974). Intergenerational equity and exhaustible resources. *The review of economic studies*, 41, 29–45.
- Stern, N. (2008). The economics of climate change. *American Economic Review*, 98(2), 1–37.
- Stern, N., Peters, S., Bakhshi, V., Bowen, A., Cameron, C., Catovsky, S., ... & Edmonson, N. (2006). Stern Review: The economics of climate change (Vol. 30, p. 2006). London: HM treasury.
- Tammi, T. (2011). Contractual preferences and moral biases: social identity and procedural fairness in the exclusion game experiment. *Constitutional Political Economy*, 22(4), 373–397.
- Tavoni, A., Dannenberg, A., Kallis, G., & Lösschel, 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.
- Thero, D. P. (1995). Rawls and environmental ethics. *Environmental Ethics*, 17(1), 93–106.
- Tremmel J. C. (2009). A theory of intergenerational justice, Earthscan.
- Tremmel, J. C. (2013). The Convention of representatives of all generations under the 'veil of ignorance.' *Constellations*, 20(3), 483–502.
- Wall, S. (2003). Just savings and the difference principle. *Philosophical Studies*, 116(1), 79–102.
- Wallack, M. (2006). Justice between generations: the limits of procedural justice, Handbook of intergenerational justice, 72–85, edited by Tremmel J. C., Edward Elgar.
- WMO (2018), Greenhouse Gas Bulletin (GHG Bulletin), No. 14. The state of greenhouse gases in the atmosphere based on global observations through 2017.

- Wolf, S., & Dron, C. (2020). The effect of an experimental veil of ignorance on intergenerational resource sharing: empirical evidence from a sequential multi-person dictator game. *Ecological Economics*, 175, 106662.
- Wolf, S., & Wagner, K. (2016). If Future Generations Had a Say: Experimental Evidence on Resource Sharing with Veto Power of a Future Generation, working paper, Department of Environmental Economics and Resource Management, University of Freiburg.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.