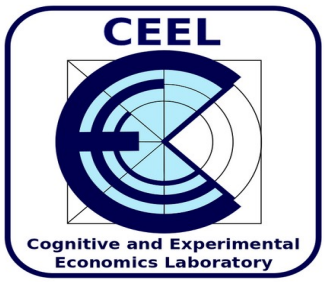


Klaudijo Klaser, Lorenzo Sacconi and Marco Faillo

Climate Change and Intergenerational Social Contract: Insights From a Laboratory Experiment in Rawlsian Perspective

CEEL Working Paper 3-20

Cognitive and Experimental
Economics Laboratory

Via Inama, 5 38122 Trento, Italy

http://www-ceel.economia.unitn.it
tel. +39.461.282313

Climate Change and Intergenerational Social Contract: Insights From a Laboratory Experiment in Rawlsian Perspective

Klaudijo Klaser[1], Lorenzo Sacconi[2] and Marco Faillo[3]

***Abstract**:* Many actions we take today will show some of their consequences in the future. Therefore future generations, although they cannot have a real voice, should be considered as direct stakeholders of some of our present decisions. As far as this intertemporal misalignment between actions and outcomes is concerned, climate change is the most evident example we have of negative externality towards the future. This paper looks at the climate change problem and the related international agreements on the reduction of greenhouse gas emission through the social contract perspective.. We apply John Rawls's veil of ignorance decision-making model within an experimental setting. In particular, we implement a sequential group dictator game where generations (groups of players) are located on a chain representing the time line. The (laboratory) veil of ignorance induces a fair ex-ante perspective regarding the distribution of resources between generations, however ex-post compliance to the agreement remains an open issue.

JEL Code: D63, D64, F64, Q54

Keywords: Experimental Economics, Climate Change, Intergenerational Allocation of Resources, Veil of Ignorance, Social Contract Theory

1 klaudijo.klaser@unitn.it, Department of Economics, University of Trento
2 lorenzo.sacconi@unimi.it, Department of Law, University of Milan
3 marco.faillo@unitn.it, Department of Economics, University of Trento

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

Edmund Burke

"The dominant reason for acting on climate change is not that it would make us better off. It is that not acting involves taking advantage of the poor, the future, and nature"

Stephen M. Gardiner

Introduction

Climate change is a threat which looms over future generations but that is triggered by some careless actions of the present one. In a technical language "[c]limate change is an instance of an externality—when one agent's activities have costs or benefits for other agents that are not reflected in the prices the first agent faces" (Clements 2015, p. 263). Thus, only the present generation can act in advance constraining its own behaviour in order to reduce negative externalities, that is the risk of bad consequences of global warming on future generations.

More importantly, climate change is a global issue and it cannot be tackled by the commitment of a minority of virtuous agents, that is environmental issues like the global warming cannot be solved through single community actions; instead they require a certain degree of international cooperation (Stern et al. 2006, p. 512). Therefore, if nations do not reach soon a widely shared agreement on how to coordinate themselves in order to limit or to avoid, through today's actions, the global warming there is the risk to harm seriously future generations.

However, "notwithstanding more than twenty years of international negotiations to establish limits, emissions of greenhouse gases continue to rise" (Gardiner et al. 2016, p. 137)[4]. In other words, even if mankind understood the dangers intrinsic in climate change long time ago (Nordhaus 1993), it is not still capable to pursue the common goal of

4 The first formal attempt to address climate change towards common solution was the United Nations Framework Convention on Climate Change in 1992.

containing the risk to harm future generations. Two are the main reasons of this constant failure of international negotiations on climate actions.

On the one hand the success of agreements concerning the reduction of global greenhouse gases (the main cause of the global warming) is essentially conditioned upon the distribution of costs between nations: by and large some of them are expected to make grater sacrifices in containing emissions, or said more explicitly, some nations are supposed to pay higher costs in pursuing the common goal (Gardiner 2011a, Gardiner et al 2016). This makes difficult any agreement, because nobody is really available to pay more than the others.

On the other hand, even if some actual agreements were formally reached (e.g. the Kyoto protocol), compliance to such agreements is known to be extremely fragile. Since reducing emissions is costly and since in the current geopolitical frame there are not institutions which can monitor and sanction defectors, single nations have a clear economic incentive to free ride[5]. Therefore, although it might be collectively rational to cooperate, from the individual (national) point of view it is rational to deviate from formal agreements.

However, without an international agreement to reduce greenhouse emissions the world will head a tragedy of commons (Hardin 1968) because the atmosphere, however big, can contain a limited quantity of greenhouse gases before they show their harmful effects on mankind. In addition, we have to take into consideration that the common-pool resource dynamic (Ostrom et al. 1994) which usually drives the appropriation of natural resources (and the resulting parallel creation of noxious waste) is amplified by a strong present bias, since action and consequences occur at different time. In other words, given the asymmetric relationship between generations, overexploitation of resources most of the times benefits the present generation at the expenses of the future ones.

For example, producing electricity through nuclear power stations we currently benefit of energy generated at a lower cost. However as for some areas the radioactive slags will constrain and will jeopardize future generations for hundreds of years. The consequences of an unlimited use of fossil fuels might be even more tragic because they involve the whole globe. The accumulation of carbon dioxide in the atmosphere generates

5 Baseically, we cannot change the payoff matrix, that is we cannot change the incentives' structure to induce or to enforce cooperation between nations because there are not global institutions which can actually do that.

the so called greenhouse effect and this, through the increase in global temperatures, will negatively condition the existence of future generations all over the world.

Therefore how can the current generation, in its own decisions, take into account in a fair way the interests of future generations which have no voice (because not existing yet) at the table discussion but which are clearly direct stakeholders of present actions? The answer is not easy at all if we take into account many specific features which distinguish allocation of resources between generations (Meyer 2016) from the circumstances which characterize the more familiar redistributive issue between contemporaries (Lamont and al. 2016 and Tremmel 2009, p. 147)[6].

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 also on the consumption (or on the utility) of future (people) generations (Solow 1974). Within this kind of functional forms a discount rate is introduced to represent the degree of concern that one generation bears for the next ones, so that the maximization of the present utility keeps balanced with the (conjectured) interests of future generations. Basically, a positive discount rate is supposed to avoid overconsumption behaviours of the present generation which might damage future generations.

However the assumption of an intergenerational perspective embodied in the utility function through a discount rate seems to be quite limiting for different reasons. For example that approach contradicts the classic pillar of the purely selfish *homo oeconomicus* who is supposed to care exclusively about his own consumption and not at all about the welfare of his offspring[7].

6 "[I]ntergenerational [distributive] justice is saddled with puzzling difficulties, such as the nonidentity problem [...], the cooperation between generations [...], motivational considerations, conceptualisation of duties and toward future generations, lack of information, uncertainty and asymmetries of power" (Gabor 2013, p. 301). To not take into consideration the rationale of deriving intertemporal norms from a purely intragenerational context (Heyd 2009, p. 177). In other words a group of existing people (the contemporaries) is supposed to derive distributive principles and practices which take into consideration interests of other non-existing groups (future generations) that formally cannot claim anything. Said otherwise, we have "to build and rationalise a problem of cooperation, duties, rights, compliance, between non-existent individuals who lived, live or will live in different moments of time" (Gabor 2013, p. 304).

7 Moreover, it is not clear why the other regarding concern is usually assumed to be a one way vertical component. In other words the following discrimination is mysterious: why would it be legitimate to assume that an economic

Furthermore, within that economic approach very much depends on the social rate of discount and on the utility functional form. However there is not unanimity on how much exactly future generations are supposed to count in our present decisions. In the environmental economics field there is still a heated debate regarding the appropriate weight (social discount rate) to assign to future generations' welfare (Nordhaus 2008, Moore et al 2004 and Stern 2008). In the same way, there is complete uncertainty about future generations' utility functions. Said otherwise, future people do not exist yet for definition and they cannot reveal their preferences (Beckerman 2006 and Parfit 1984), therefore we cannot really know what is better for them (Barry 1977).

It is within this (considered inadequate) theoretical framework that the social contract theory can provide a useful and innovative tool to deal with climate change and more generally with distribution of resources between generations.

John Ralws (1999) was the pioneer in extending in a structured way the social contract model to the allocative problem between generations. Within his theory, Rawls claims that the set of currently existing people (therefore not all the generations of the history), instrumentally rational and free of any other regarding preference, have to design the principles to regulate the intergenerational distribution of resources (Rawls 1999, pp. 118-123). Preventing then, by means of a veil of ignorance, the current generation from knowing the specific moment of the history it belongs should induce the (present) parties to design a fair principle for the allocation of resources through the human history.

The most interesting feature of the Rawlsian intragenerational setting for intergenerational principles is exactly the designed decision-making model. In particular in Rawls's intergenerational theory individuals taking part in the agreement are only contemporaries and they have to choose an allocative configuration which does not affect exclusively themselves. Instead the present generation is called to evaluate distributive principles that will produce effects also on third parties (future generations) who formally cannot take part in the contract, cannot make demands, cannot make objections, cannot

agent cares about her children, but not about the welfare of her siblings or her friends, or even about her ancestors (like parents or grandparents)? The first one is usually considered a natural fact, the other comparable social preferences are instead seen as exceptions.

threat and cannot punish the actual decision maker: this is basically the modern climate change issue[8], which is the result of a substantial dictatorship of the present.

A further relevant feature of a Rawlsian approach to the climate change issue is then the ex-post compliance to the social contract. Indeed the agreement behind the veil of ignorance is not conceived by Rawls as binding. In other words, although the distributive principles are the outcome of a formally fair procedure, once the veil is dropped a dictatorship of the present generation over the future ones remains a concrete possibility. This is a strong analogy with what was previously shown: compliance to international agreements on emissions reduction is a big issue, because they are not enforceable.

In the experimental field compliance to non binding intragenerational distributive principles was explained with the Rawlisan idea of the sense of justice (Degli Antoni et. al 2016, Faillo et al. 2008, Faillo et al. 2014, Rawls 1963, Rawls 1999, Sacconi and Faillo 2005, Sacconi and Faillo 2010, Sacconi et al. 2011 and Tammi 2011). Therefore we can amplify the same concept of the sense of justice to the intergenerational context. In particular, we can extend its validity if we verify that a different decision-making framework (with the agreement that does not formally and substantially include all the stakeholders), concerning a slightly different distributive problem (resources have to be distributed not within group but between groups), leads to the same conclusions on the sense of justice.

Within the general framework described so far, we apply Rawls's intergenerational social contract model within an experimental setting with the aim to provide some insights concerning the modern climate change issue.

Next Sections are then organized as follows.

Section 1 introduces Rawls's social contract theory (Rawls 1999) focusing on its intergenerational extension. The aim is to provide the theoretical background for the analysis of distribution of resources between generation in a Rawlsian perspective. Although

8 The social contract on intergenerational principles is very different from the standard distributive issue where a set of (contemporary) individuals has to agree on the way to divide resources among themselves: whether I have to agree with you on how to split $10 between us or whether I have to agree with you on how to split $10 between ourselves and a third person who has absolutely no voice on the issue are two extremely different decision-making situations. In the latter case individuals involved in the contract are supposed to take in consideration some people who remain outside of the contract itself.

the slant given by John Rawls to his intergenerational social contract theory (Rawls 1999, pp. 251-267 and Rawls 2001, pp. 158-160) requires some prudential clarifications (see Appendix A), his idea of an intragenerational agreement behind a veil of ignorance (Rawls 1999, pp. 118-123) is considered an adequate model for inquiring the modern climate change issue.

Section 2 describes the experimental design which captures the main features of the modern climate change problem in a Rawlsian perspective and it provides the predictive hypothesis. The game is structured as a group dictator game (Kahneman et al. 1986) played sequentially (Bahr et al. 2007 and Casol et al. 1998) and it is run in two distinct conditions, with and without a preliminary voting stage simulating an agreement behind a veil of ignorance.

Section 3 analyses the data of the experiment.

Appendix A deepens the Rawlsian social contract theory extended to the intergenerational issue, while Appendix B contains the instruction provided to participants and read aloud during the experimental sessions.

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

Rawls's ethical system (Rawls 1999) is designed to identify the main principles which should lead the human society and its institutions, particularly with regard to the division of benefits generated by cooperation between individuals (Rawls 1999, p. 4). With his theory Rawls establishes a procedure inspired by the social contract tradition, that is the principles are the outcome of an agreement between those individuals involved in the cooperative scheme.

Within the Rawlsian decision-making procedure the impartiality in the choice of the principles is guaranteed by a veil of ignorance (Rawls 1999, pp. 118-123). This is a tool which excludes the access to any particular information to those parties who take part in the agreement.

In this sense "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. Nor, again, does anyone know his conception of the

good, the particulars of his rational plan of life, or even the special features of his psychology such as his aversion to risk or liability to optimism or pessimism". Furthermore it is important to highlight how according to Rawls "persons in the original position have no information as to which generation they belong" (Rawls 1999, p 118).

"In this manner the veil of ignorance is arrived at in a natural way", since it is excluded "the knowledge of those contingencies which sets men at odds and allows them to be guided by their prejudices" (Rawls 1999, p. 17). In other words behind the veil of ignorance no one can take advantage of personal contingencies to design principles (norms or institutions) which might favour her own particular position. Thus, in the original position everybody is equally represented since everybody has to choose in the same situation of perfect (mis)informational symmetry and the involved parties reach an agreement only on the basis of impartial and general considerations (Rawls 1999, pp. 118-123).

Although the distribution of resources between generations is particularly challenging and it "subjects any ethical theory to severe if not impossible tests" (Rawls 1999, p. 251) John Rawls (1999 and 2001) does not fail to extend his contractarian system to contemplate this relevant topic. Indeed, Rawls is aware how the account of his social contract theory "would be incomplete without some discussion of this important matter" (Rawls, 1999, p. 251).

When Rawls moves from the intragenerational context to the analysis of principles that are supposed to regulate the allocation of resources between generations he adds a reasonable specification concerning the decision-making procedure described so far. In particular, Rawls specifies that although people are deprived of the information concerning the generation they belong (Rawls 1999, p. 118 and p. 254), 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, people involved in the agreement, even though focused on intergenerational principles, belong to the same generation (and they know it as general fact).

In this way Rawls substantially constrains his intergenerational decision-making model to the physically existing people (Dierksmeier 2006, p. 74). Indeed, in his opinion it would be unrealistic to conceive an agreement (although hypothetical) which gathers

together all the possible generations of the human history: this all-inclusive approach would stretch imagination too much, that is it would require a too high level of abstraction[9] (Attas 2009 pp. 195-7, Rawls 1999 p. 120, Rawls 2001, p. 160 and Tremmel 2009, p. 156).

Notwithstanding the specification concerning the contemporaneity of the parties (see Appendix A), Rawls's social contract theory seems to be promising for dealing with the intergenerational allocation of resources (Tremmel 2013, p. 484), because thanks to the veil of ignorance the present generation substantially looses its privileged position (dictatorial powers) towards future generations. Thus the veil of ignorance guarantees that the parties involved in the agreement, despite being and knowing to be contemporary, are encouraged to propose impartial principles for the division of resources between generations[10].

However, in Rawls's opinion the standard difference principle, which requires to maximize the expectations of the worst-off (Rawls 1999, p. 56, p. 69, and p. 72), is not a suitable tool to deal with redistribution of resources between generations because it apparently produces some undesirable consequences (Gardiner 2009, Rawls 1999, pp. 253-255 and Appendix A). Therefore the issue concerning the intergenerational allocation of resources "must be treated in some other manner" (Rawls 1999, p. 254). Thus Rawls proposes the just saving principle as normative rule to to regulate distribution of resources over time: "in following a just savings principle, each generation makes a contribution to those coming later and receives from its predecessors" (Rawls 1999, p. 254).

Nevertheless Rawls does not describe in detail the peculiar features of the just saving principle (like for example providing a specific saving rate or a schedule of rates). Instead he limits himself to sketch some general ethical restriction which the contractual parties should take in account in defining the saving path (Rawls 1999, pp. 255-6). However, more details about the derivation and the configuration of the just saving principle are provided in Appendix A.

---

9 Indeed Rawls's model is to be "understood as a purely hypothetical situation" (Rawls 1999, p. 11). In other words, Rawls's original position coincides with the adoption of a particular perspective, and therefore the agreement is conceived as a simple mental experiment.

10 According to Rawls, even though it is formally the present cohort to decide about the allocation of resources between generations, the veil of ignorance procedure induces the parties to take into appropriate consideration also future generations and to choose "a path over time which treats all generations justly during the whole course of a society's history" (Rawls 1999, p. 257)

For the purposes of the experiment and of its interpretation it is sufficient to highlight the working mechanism of the just saving principle. Basically, 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. This without the actual participation of the latter in the initial agreement. Therefore, with Rawls's social contract model (and 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 due consideration individuals substantially left outside of the agreement (future generations).

2. Experimental design and predictive hypothesis

Many different theoretical approaches were proposed to address the environmental issue in a Rawlsian perspective, going from considering health and environment as social primary goods and including animals in original position (Gardiner 2011b and Thero 1995), to running a third level original position (Clements 2015). In the same way many experimental works aimed to test empirically the assumptions or the conclusions of Rawls's social contract theory (Gaertner and Schokkaert 2012). Nevertheless, there were no concrete attempts which explicitly tried to merge the two fields. To the best of our knowledge, there was only one single attempt to test Rawls's intergenerational theory within an experimental setting (Wolf and Dron 2015).

Wolf and Dron's design is very intuitive. A common endowment is provided to five people. Single players are then randomly assigned to a position within a sequential dictator game (Bahr and Requate 2007 and Cason and Mui 1998).

Starting from the player occupying the first position, participants sequentially enter the dictator role and they are asked to claim a share of the common endowment for themselves[11] until either the fourth player takes a decision (the fifth player substantially becomes a dummy player) or the common endowment is exhausted. Indeed, in order to represent realistically the "dictatorship of the present" issue, every player who enters the

11 The share players decide to withdraw from the common endowment during their turn constitutes their final payment.

dictator role is allowed to claim the 100% of the (remaining) endowment, with the consequence that nothing would be left to the following players.

The game's underling idea is that each single player in the sequence formally represents a (non-overlapping) generation, because each person: *a)* (except the first one) is subject to the consequences of the decisions taken by all previous generations; *b)* with her or his own decision can influence only the welfare of the following generation(s).

In one of the proposed treatments Wolf and Dron introduce a preliminary stage where players are asked agree on a rule to share the common endowment between the five generations in the sequence. However, they are asked to do that behind a veil of ignorance, that is before knowing the position they will occupy in the actual sequential dictator game. After 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.

However, Wolf and Dron's (2015) attempt is to be considered unsuccessful. First of all the veiled agreement did not produce a (significant) more equal distribution between the five players compared to the baseline treatment where there was not any kind of agreement before players entered the sequential dictator game. Despite players had the chance to discussed about the issue, players in privileged position profited anyway of their contingency (Wolf and Wagner 2016).

Even more importantly, their experimental design is to be considered inconsistent with regards to Rawls's theory interpreted strictly. Indeed in Wolf and Dron's experiment all the generations are put behind the same veil of ignorance, as generations could reach an intergenerational agreement (Anderson 2013). Instead John Rawls is really clear and careful to specify how the veil of ignorance for intergenerational principles is intragenerational, that is only contemporaries, generation by generation, are involved in the deliberative process behind the veil (Rawls 1999 pp. 118 -121). Thus, since Wolf and Dron's design represents generations with single individuals, no formal intragenerational agreements are possible[12].

Notwithstanding the imprecise design and the discouraging outcome, Wolf and Dron's provide a basis to inquire Rawls's intergenerational theory, mainly because it follows

12 It is true that according to Rawls the agreement behind the veil of ignorance is a mere mental experiment (see footnote number 6), so a formal agreement is not necessary in order to derive the principles of justice. At the same time, in a laboratory experiment it seem to be too ambitious to simulate an agreement with the unilateral decision of one single person.

a widespread practice in the economic experimental literature, that is to simulate generations assigning players to different positions in a sequential game[13].

Thus, in order to design our experiment we started from Wolf and Dron's sequential dictator game and we improved it taking into account the further specification made by John Rawls concerning the nature of the impartial agreement for intergenerational principles. In particular we focus of the nature of the agreement: 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), all of them strictly belong to the same generation, that is the impartial agreement is intragenerational and involves only contemporaries.

Thus, in our particular design generations are constituted by groups of three players[14]. Groups are then randomly assigned to a position on chains (sequences) of different lengths[15]. Essentially, every group is meant to represent a (non-overlapping) generation of contemporaries. Starting from the first generation players are asked to play a group sequential dictator game.

During the game nobody can know the total length of her own chain because it is not communicated[16]. However players can deduce how many generations exist (how many groups play) before their own (in case) enters the dictator role since all chains start with the generation number one. For example, if a group is assigned to the generation number three, players of that group do not know how many generations there might are after theirs, but they know for sure that other two groups have to play before they can possibly take any decision.

13 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). Usually in this kind of intergenerational experiments there is not any strategic interaction between players belonging to different positions in the sequence, because later generations cannot directly influence the payoff of the previous ones (while the opposite is true).

14 According to the introductory framework, the single parties in the experiment should be considered representatives of states , like in the Law of People (Rawls 2001a, p. 10).

15 The shortest chain was made of one group (generation), the longest one of a sequence of five groups (generations).

16 This is a standard practice in experiments of this kind (see for example Fischer et al. 2004 or Hauser et al. 2014). This hidden information basically avoids that generations think about the last one as a pure dummy which is not supposed to take any decision.

The decision-making task for the dictator group is then designed as follows (for further details it is possible to consult Appendix B, which contains the instructions provided to the experimental subjects).

The first group (generation) of each chain has at its disposal a common endowment of €21. Each of the three participants of the group individually has to decide how much money to withdraw from the common pool, choosing an integer value between €0 and €7[17]. The amount each player claims for himself in this stage constitutes her individual final payment. After a player makes a choice and before he is revealed the outcome of the group he is asked to guess, through an incentivized structure,[18] the decisions taken by the other two players of his group.

After all the 3 players decide how much money to withdraw from the common endowment, one of the two following scenarios occurs:

- if the common pool is left with at leas €6 in total, the chain continues and the next group in the sequence enter the decision-making phase becoming the dictator group. The common endowment is refilled up to the initial value of €21[19] and the new generation faces the same identical decision-making problem described so far[20];

17 It is important to remark that given the structure of the game players belonging to the same group are endowed with symmetric capabilities, that between players within a single generation there are neither formal nor substantial differences. In the game differences between players are exclusively relevant with regard to the group position in the chain. Therefore within this design one of the two problems linked to the unsuccessful international agreements on climate actions, that is the distribution of individual costs, is basically put aside. Indeed within a situation of perfectly symmetric roles there are not formal reasons to distribute costs unequally. However, this simplification does not make the experiment less useful in order to solve the climate change issue: if nations were not able to agree even in a situation of symmetry, a fortiori we could not expect a widely shared agreement in the case of asymmetric costs and benefits. Therefore our experiment constitutes an important step in understanding international agreements on the reduction of greenhouse gases.

18 Players with the best guess were rewarded with €2 extra. Given the symmetry of the roles within a group, 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). See Appendix B for further details.

19 The technology is certain and identical for every group in the sequence.

20 The choice of refiling the pool up to the same initial level, despite it might appear unrealistic, it was made to facilitate the agreement framework in the veil treatment. Indeed, having one unique level of endowment which is certain ensures that every group reaches the agreement in the same structural conditions. This allows also an easy interpretation of the compliance task, because this one will not depend on the actions taken by previous generations. Moreover, giving an agreement, it is always possible to comply with it, because the agreement and the decision-making frameworks coincide. However,, the practice of providing the same endowment to every generation is not

- if in the common pool players leave a total of €5 or less, the common pool is emptied and the chain breaks up, with the consequences that all the following generations cannot take any decision and do not get paid.

An experimental session lasts up to the point that all chains either get to their natural end or break up.

The minimum material threshold of €6 has a clear interpretation. It simulates the threat embodied in climate change. If the present generation overexploits the environment and does not constrain itself in consuming some available resources which can increase its own welfare, it does that at the expense of all future generations. On the contrary, if players of the group called to take a decision (present generation) coordinate for not overexploiting the environment and for leaving a minimum amount of resources for the next generation, the latter can enjoy the same opportunities as the former.

The experimental design tries to mimic as much as possible the following issue: if we want to take into account the interests of future generations in a fair way, avoiding as much as possible the global warming consequences, we must coordinate (constrain ourselves) to reduce today the consumption of fossil fuels. Instead, if active players (those who enter the dictator role time after time) do not take into sufficient consideration the interests of following groups, the former can seriously harm all future generations.

The veil treatment, which adopts the Rawlsian insights to address the concern for future generations, adds to the baseline treatment described so far a preliminary stage where the three players of every single group have to reach an (intragenerational) internal agreement in order to enter the (intergenerational) group dictator game. In particular, at the beginning of the experimental session every group is asked to unanimously agree on one the two following rules which are meant to deal with the common endowment:

- *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. This rule is meant to represent scheme consistent with Rawls's just saving principle.

new in the experimental literature (e.g. Hauser et al. 2014).

- *Interruption of the chain*: having the possibility to do it each participant of my group should withdraw from the common more than €5, even if that means interrupting the chain.

Every group is asked to reach a unanimous agreement behind a veil of ignorance, that is before being assigned to a position in a chain. Thus, consistently with Rawls's setting, while groups of contemporaries vote for a principle aimed to manage the appropriation of common resources, they do not know the generation (position) they belong in the history (in the chain).

In both treatments, at the end of the experiment a general socio-demographic questionnaire was provided.

Before moving to the predictive hypothesis, two further features of the veil treatment deserve attention.

In the first instance, we need to clarify the interpretation we give either about those groups who might not reach an agreement at the beginning of the veil treatment or about those groups who formally reach an agreement, but cannot *de facto* play any game because a previous group of players left less than €6 in the common pool, breaking in this way the chain up.

With regard to the former case, the interpretation seems to be quite intuitive. Groups (generations) that do not reach an agreement end up with living in the so called "state of nature". In other words generations who do not agree to enter a society built on the cooperative attitude and on mutual advantage enter anyway the intergenerational chain (the history), but they do not put themselves in the minimum essential condition to exploit the available resources (the common endowment of €21). We have to imagine a situation where fossil fuels are fully available in nature and ready to be exploited. Nevertheless, the generation of people who did not reach a preliminary agreement can only look at those resources without being able to "touch" them, because the concerned parties did not agree to cooperate in order to organize their extrapolation. Thus people who did not reach an agreement come to the existence but they live in poverty because they cannot exploit the available resources.

As far as the latter case is concerned, the interpretation seems to be even more straightforward. Even if a group agreed to enter a cooperative society, they end up with

living in poor conditions as well. However this time it is not for their own (missing) willingness too coordinate and too cooperate, but because some of the previous generations did not leave enough resources to allow their society to be wealthy. In this case we have to imagine a situation where the unlucky society, despite having reached an agreement, observes that fossil fuels (the common endowment of €21) are not physically available because previous generations overexploited the nature. In this perspective the unlucky generation basically pays the consequences of the global warming generated by the previous unconstrained behaviours.

Thus, the two mentioned situations are identical about the substantial material consequences on existing generations: in none of the cases they can enjoy the common endowment. However, they are induced by different formal causes: they do not coordinate in the first case, while they suffer the decisions of other groups in the second case.

Second, it is important to remember how the agreement reached behind the veil of ignorance is not conceived by Rawls as binding. Said otherwise, after the veil is dropped and groups are assigned to a chain and to a position, the outcome of the agreement is not automatically implemented (like it was did in other intergenerational experiments, e.g. Hauser et al. 2014[21]).

Thus, in our design, generations that are called to take a decision are not constrained by any external enforcement mechanism to apply the outcome of the agreement reached behind the veil of ignorance. This implies that in the veil treatment the sequential dictator game exactly replicates the baseline treatment, and compliance to the agreement is left to an individual choice. Again, this is a realistic structure since in the real world we have no

21 The mentioned practice is considered unrealistic. Even though the authors justify the binding vote as a good proxy for informal institutions which usually enforce cooperative attitudes (like punishments or rewards), those enforcement mechanisms work only when there are repeated interactions among the same subjects, so that paying a cost now (punishing) can generate long-term benefits. Indeed, this kind of institutions cannot be as much effective in one-shot games as in repeated games. In their Intergenerational Good Game there are no rational reasons (except maybe spitefulness) to punish another player who did not comply with the approved rule since I will not interact with him in a next round. I could only loose by materially punishing somebody else belonging to my group (generation). Therefore, the pretension to assimilate a binding vote to an informal institution which can enforce cooperation has to be considered inadequate, at least for a context simulating an intergenerational game played sequentially.

formal institutions which can substantially constrain the present generation to care about the future ones, even if formal intergenerational norms are the outcome of a fair procedure.

Given the theoretical framework and the experimental design described so far we can proceed with formulating the predictive hypothesis of our game. Our first hypothesis regards the baseline treatment and it follows from standard economic assumptions[22]. Without any other formal element in the game, the sub-perfect equilibrium is represented by the triple (€7, €7, €7) for every (including the first) 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. Therefore chains will not continue after the first generation, because behaviours of the first generation undermine the entire scheme of indirect cooperation over time. Thus,

H1: in the baseline treatment the generations number 1 will mostly break all the chains up

Our second hypothesis directly follows from the Rawlsian theory. As we have seen, according to Rawls, groups behind the veil of ignorance should agree "on a path over time which treats all generations justly during the whole course of a society's history" (Rawls 1999, p. 257). More precisely, players should agree on a just saving principle, according to which "each generation makes a contribution to those coming later and receives from its predecessors" (Rawls 1999, p. 254). Therefore to guarantee a positive saving to each generation that allows the chains to continue, most of the groups are expected to agree on a rule which somehow constrains a pure self-interest behaviour. In short, players should agree that each individual is supposed to withdraw maximum €5 from the common endowment.

H2: in the veil treatment, because of the impartial perspective, during the voting phase groups will mostly agree on the rule representing the Rawlsian just saving principle which guarantees the continuation of the chain

22 Except for special cases (Bardsley 2008, Cherry et al. 2002 and List 2007) threshold established in our experiment (6/21) mirrors exactly the average amount of money left by dictators (28.5%) to dummy players (Engel 2011). However, the impossibility to coordinate represents a non.-indifferent obstacle to allow chains to continue, since one purely selfish dictator can nullify the effort of the other two.

The third assumption follows as much from Rawls's social contract theory as from the experimental literature based on it: although the agreement behind the veil of ignorance is not conceived as binding, individual compliance to the chosen principle is expected to be high even in those cases where players agreed on a counter-maximizing rule. Whereas for the standard economic theory, since the agreement does not introduce formal constrains, every individual in the decision-making phase should follow his purely selfish impulses claiming €7 regardless to the chosen principle behind the veil of ignorance,.

The so called exclusion game (Degli Antoni et. al 2016, Faillo et al. 2008, Faillo et al. 2014, Sacconi and Faillo 2005, Sacconi and Faillo 2010 and Tammi 2011) inquired from an experimental point of view the Rawlisan concept of the sense of justice (Rawls 1999). In particular, the exclusion game is a one-shot resource allocation game contemplating also a preliminary voting stage carried out behind the veil of ignorance: during the voting phase parties are prevented from knowing their role in the actual game (dictator or dummy).

In the game the agreement concerning the sharing rule is not geared towards being binding, therefore players who are assigned the dictator roles are supposed to pursue their own interest regardless the rule agreed in the voting stage. This is a clear analogy with the present generation in our intergenerational experiment. However the experimental evidence of the exclusion game shows how the (unconstrained) ex-post compliance with the ex-ante chosen distributive norms is unexpectedly high even in those cases where groups agreed on an egalitarian (counter-maximizing) distributive rule[23]. Therefore we expect that

H3: in the veil treatment individuals will comply with the intragenerational agreement

The last hypothesis becomes a logical sum of the previous two: if groups agree on an intergenerational sustainable behaviour (just saving principle), and if they comply with the chosen norm, chains will continue up to the last generation.

23 However, we have to keep in mind the dissimilarity between the exclusion game and our intergenerational agreement: the agreement of the former includes all the interested parties (including the dummy player), while the latter in the voting phase leaves out some of the direct stakeholders. This might have an impact on reciprocal conformity and therefore on compliance.

H4: compared to the baseline treatment, in the veil treatment a significantly higher number of chains will continue until their natural end

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

H4a: the proportion of people claiming an amount of €5 or less will be higher in the veil treatment than in the baseline condition

and

H4b: the average individual claim will be lower in the veil condition than in the baseline treatment

3. Data analysis and comment

All the experimental sessions took place in the Computable and Experimental Economics Laboratory (CEEL) of the University of Trento. They were run using the free software for economic experiments zTree (Fischbacher 2007). All participants who took part in the experiment showed up after a public call.

In the experimental laboratory participants were randomly assigned to a computer terminal. All the emplacements were isolated by separation walls to avoid communication. 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 at all clear, and 5 = very clear). Before the actual experiment could start, in the baseline (veil) treatment 4 (6) control questions about the structure of the game were asked.

The experiment involved a total of 141 participants (60 in the veil treatment and 81 in the baseline treatment). On average participants were 22, 54% of them were females and 46% of the total participants were enrolled in programs related to the economic discipline, the rest in other fields going from humanities to natural sciences. 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 maximum 45 minutes.

In the baseline treatment we run 4 sessions with a total of 8 chains (Table 1)

Table 1 – Distribution of chains per session in the baseline treatment

Session	Number of chains	Chains' length (n. generations)
1	2	2 and 5
2	2	2 and 4
3	2	2 and 5
4	2	3 and 4

In the baseline treatment 6 chains out of 8 (that is 75%) broke up after the choices made by players belonging to the first generation. In the two remaining chains (session 1 and 3), despite having only two generations, the second generations substantially behaved such that they would have broken their chains up. Thus out of 81 participants only 30 (37%) of them played in the active role of the game taking an actual decision. These first data are sufficient to support ***H1***: in the baseline treatment chains (mostly) brake up after the first generation.

In the baseline treatment active players withdraw €5.30 on average from the common endowment believing that the other two players in the group would have claimed €5.20 on average. These last data show an interesting empirical regularity which was not taken into account by the predictive hypothesis: generations in the baseline treatment waste resources[24], because on average they left in the pool 1.70€, an amount of money which in the game is basically destroyed since it is not distributed to anybody. Indeed, despite left by the active players with the hope to contribute to the next generation, that amount was not enough to allow chains to continue.

Seen from another point of view, we can claim that some players are altruistic individuals. Indeed, they renounce to consume (to withdraw) a part of their individual endowment without any possibility of being reciprocated by future generations. More specifically, they are intergenerationally altruistic, because they take into account the interests of possible future generations at expenses of their own material payoff. However

24 In general all active groups left in the common pool at least €2. A total of €36 was left to the experimenter.

their good purposes are nullified by the actions of a minority of players who do not do their part in the fair management of the common resource.

With regards to the preliminary voting stage of the veil treatment, all groups reached a unanimous agreement on the rule by the second round (out of six available). 18 out of 20 groups agreed on the rule "continuation of the chain". Therefore our ***H2*** is supported by the data too, because most of the participants (54 people, representing the 90%) voted for the just saving principle.

In the veil treatment there was a total of 6 chains divided in 3 sessions (Table 2)

Table 2 – Distribution of chains per session in the veil treatment

Session	Number of chains	Chains' length (n. generations)
1	2	2 and 5
2	2	2 and 4
3	2	3 and 4

After the agreement 80% of the participants who took a decision complied with the rule chosen behind the (laboratory) veil. Thus, 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.

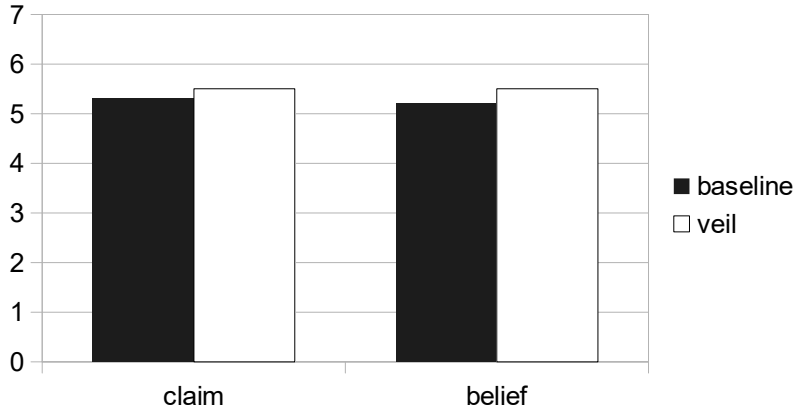
Nevertheless, although participants mostly choose the just saving principle and they mostly complied with it, in the veil treatment no one chain continued after the second generation. Therefore, ***H4*** is rejected by the empirical data, because no one chain got to its natural end as predicted.

That ***H4*** is not supported follows also from the rejection of the sub-hypothesis. ***H4a*** is rejected since between treatments the proportion of 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 veil treatment players claimed for themselves an average share of the common endowment not statistically different from the baseline treatment (two tails Welch *t* of student test is t= -0.78, with a corresponding *p-value* of 0.44), withdrawing on average €5.50[25] (Chart 1).

25 Braking the chains up left to the experimenter a total of €21. Therefore in the veil treatment active players wasted on average €1.40, making the saving more efficient compared to the baseline.

Chart 1 – Individual average claim and average belief (€) per treatment



The undifferentiated behaviour which rejects ***H4b*** and then ***H4*** can be highlighted also with a simple linear model (1) where we regress the individual claim on the *treatment* (dummy variable, where the baseline assumes the value 0), on the *generation* (position in the chain), on the *average belief* (in €) plus a series of control demographical factors, clustering the standard error per single groups (Table 3).

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

Table 3 – Determinants of individual claiming[a]

treatment	0.18 (0.29)
generation	0.53 (0.267) *
average belief	0.48 (113) **
clearness instruction	0.35 (134) **
n. previous experiments	0.04 (0.005) **
age	- 0.07 (0.080)
gender	- 0.39 (0.264)
nationality	0.75 (0.307) **
field of studies	- 0.4 (0.218)
years of studies	0.11 (0.108)
risk attitude	0.04 (0.037)
family yearly income	0.24 (0.110) **
general economic situation	0.50 (0.213) **
* significant at 5%, ** significant at 1%	
Number of observations 48[b]	
Adjusted R-squared 0.6013	
F-statistic 6.519e-07	

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 foreigner), *field of studies* and *year studies* refer to the undertaken studies (the first one is a dummy variable, where the value 0 corresponds to economic disciplines, whereas 0 to any other field, 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 the yearly family income and the current economic status (both are discrete variables and range from 1, very low/bad, to 5, very high/excellent).

b The number of observations is lower than stated so far because we could not collect the demographic statements of three participants in the baseline treatment, so the model drops those data. However, those three observations did not influence the outcome since the model considering only the three fundamental variables (*treatment*, *average belief* and *generation*) do not differ from the one presented here

From the econometric estimation we can see that there is not any statistically significant difference in the average claim between the two treatments. Among the main predictors, only the average beliefs on the others' withdrawal and the generation had a strong and significant effect on the individual choice. In other words, a higher expectation of €1 on the others' withdrawal and belonging to the second generation (the last generation who actively played) increased the average claim of about €0.5.

Even though the veil treatment did not show the expected effect on the individual claim and therefore on the chains survival, there is a further interesting comparison between the two conditions we may look at. In particular, in the veil treatment only 50% were interrupted by the first generation (compared to the 75% of the baseline treatment). Thus, thanks to the agreement 45% (27 out of 60) of participants who took part in the veil treatment had the possibility to make an active choice, while that ratio lowers to 37% (30 out of 81 subjects) with regards to the baseline treatment. However, this results are not statistically significant (*chi square* test 0.29 on the proportion of participants who actively chose).

Conclusions

With our laboratory experiment relying on Rawls's intergenerational social contract theory and some of its experimental evidence we tried to address the modern issue regarding climate change actions and in particular climate change agreements. We left aside the problem concerning the distribution of costs between nations assuming symmetrical situations between participants belonging to the same generation in the game. Instead, we focused on the pure intergenerational problem, trying to see if the Rawlsian theory could help to structure a fair intragenerational agreement for the intergenerational distribution of resources.

The experimental results showed that a laboratory veil of ignorance induces people to reach an ex-ante fair agreement concerning the management of common resources and consistent with the Rawlsian just saving principle. At the same time compliance to that principle, despite being high, was not sufficient to allow chains to survive significantly longer than in the baseline treatment, where no agreements were possible.

This specific result first of all highlights what is considered a limit of our experimental design: even though the average compliance rate (about 80%) is high and consistent with the previous empirical evidence provided by the exclusion game, this percentage seems to be low compared to the 100% of compliance (to the just saving principle) required by the game to allow chains to continue.

In other words, in our experiment a sustainable society (a chain) can continue only if compliance to the just saving principle is total. This is certainly an amazing result from a theoretical point of view, but it is too demanding to expect perfect voluntary compliance, as much in an experiment as in the real world.

However, our non positive result requires also a deeper reflection which goes beyond the compliance threshold. If we look closer at the distribution of claims in the two treatments (Table 4) we can better understand why our institutional mechanism of the veil of ignorance did not work as predicted. In the Table 4 within group claims were ordered in an ascending triples order moving from the left column (player with a low claim) to the right (player with a medium and high claim). We highlighted in grey the participants who withdrew €6 or more.

Table 4 – Distribution of individual claims (average belief) per treatment and group

BASELINE TREATMENT				**VEIL TREATMENT**		
Player low	**Player med.**	**Player high**		**Player low**	**Player med.**	**Player high**
4 (4.5)	4 4.5)	4 (4.5)		5 (5)	5 (5)	5 (5)
4 (4)	5 (2.5)	7 (6)		5 (5)	5 (5)	7 (5)
5 (5)	7 (6.5)	7 (7)		4 (4.5)	5 (5)	5 (5)
4 (5)	5 (5)	7 (6)		5 (5)	5 (5)	7 (5)
4 (5)	5 (4.5)	7 (6)		5 (5)	5 (7)	7 (7)
5 (5)	5 (5)	5 (7)		4 (5.5)	5 (5)	7 (5)
5 (5)	5 (5)	6 (7)		5 (5)	5 (5)	5 (5)
4 (4)	5 (5)	7 (7)		5 (5)	5 (6.5)	7 (7)
5 (2)	5 (5.5)	6 (6)		7 (7)	7 (7)	7 (7)
5 (5)	5 (5)	7 (7)				

As also pointed out earlier in the paper, given the structure of our game one high claim (€6 or €7) in a group is sufficient to undermine the entire scheme of intragenerational cooperation necessary for the intergenerational continuation of the chain. From the Table 4 we can clearly

observe how the agreement phase of the veil treatment did not produce any effect on the participant who withdraws a high amount from the common pool. In other words, there is a share of subjects who is somehow indifferent to the impartial agreement because the last one cannot change their psychological equilibrium.

Seen from another perspective, it seems that the sense of justice to the agreement is not triggered when we agree with a set of people X to undertake a specific action towards a set of people Y (different from X). As highlighted in the Introduction, within the Rawlsian theory a generation of contemporaries is called to evaluate distributive principles that will produce effects on third parties (future generations) who formally cannot take part in the contract. This is to say the ex-post actions of mutual advantage (savings) are not directed towards the same set of people (contemporaries) we agreed the principles with.

In other words the Rawlisan sense of justice based on mutual expectation of compliance which is supposed to be the glue which ensures general compliance does not enter into play when we have impartial agreements concerning some subjects who are external to the agreement itself. Indeed, after the veil is dropped, we technically do not enter in a mutually advantageous relationship with our contemporaries. Certainly we have a common goal with them, but we cannot say we mutually benefit each other.

Therefore, when this ex-post mutually beneficial situation is not taking place like in our game we cannot expect participants to develop sense of justice towards the agreement reached with their fellows in the group of contemporaries. And without the possibility of sense of justice there cannot be compliance to the (however intergenerationally impartial) agreement. And without compliance we cannot expect chains to continue.

The big issue on how to structure international agreements on climate actions keeps open.

Appendix A: Rawlsian intergenerational justice and derivation of the just saving principle

The starting point of Rawls's reflection about allocation of resources between generations is an extension of the main hypothesis of his social contract theory, which portrays the human society as a venture for the mutual advantage (Rawls 1999, p. 4). Thus Rawls assumes that "life of a people is conceived as a scheme of cooperation *spread out in historical time*" (Rawls 1999, p. 257, *emphasis added* and Rawls 2001). Therefore, according to Rawls, it is necessary to agree "on a path over time which treats all generations justly during the whole course of a society's history" (Rawls 1999, p. 257).

Since for Rawls "persons in different generations have duties and obligations to one another just as contemporaries do" (Rawls 1999, p. 258) and since according to him justice between generations "is to be governed by the same conception of justice that regulates the cooperation of contemporaries" (Rawls 1999, p. 257) it might seem to be reasonable to extend the standard (intragenerational) principles of justice (Rawls 1999, pp. 47-101)[26] over the time dimension.

Even more relevant for the present discussion about redistribution of resources, given the just mentioned similarities, it might be intuitive to adopt the canonical formulation of the so called difference principle[27] to regulate allocation of resources between generations. After all Rawls himself explicitly claims how the "appropriate expectation in applying the difference principle is that of the long-term prospects of the least favored extending over future generations" (Rawls 1999, p. 252). Thus it seems to be that the difference principle, when fully applied, has to take into consideration and to operate on two dimensions, space and time.

However, almost contradicting his own claims, Rawls remarks through many passages how the difference principle's prescriptions have to be realized exclusively within an intragenerational context. Indeed, in Rawls's opinion, the difference principle is

26 "The first principle simply requires that certain sorts of rules, those defining basic liberties, apply to everyone equally and that they allow the most extensive liberty compatible with a like liberty for all" (Rawls 1999, p. 56), meanwhile the second principle of justice, the so called difference principle, prescribes to "maximize the expectations of the least favored position" (Rawls 1999, p. 69).

27 For a complete presentation of the difference principle see (Rawls 1999, pp. 52-65, pp.130-9 and pp. 153-160).

inadequate to discipline the allocation of resources between generations because of its undesirable consequences: "for when the difference principle is applied to the question of saving over generations, it entails either no saving at all or not enough saving to improve social circumstances sufficiently so that all the equal liberties can be effectively exercised" (Rawls 1999, pp. 253-254)

In other words, “since the persons in the original position know that they are contemporaries [...] they can favor their generation by refusing to make any sacrifices at all” for the others (Rawls 1999, p. 121). Rawls tries to suggest that since the parties involved in the agreement (who are contemporaries) are instrumentally rational and they desire to maximize first of all their own expectations[28], it is not legitimate to expect, by the generation involved in the agreement, any renounce of resources which could benefit (the least advantaged) people in another generation (Rawls 1999, pp. 254-255, Attas 2009, p. 190 and Buchanan 1987, p. 250). Thus this way of reasoning is not compatible with the most deep meaning of the difference principle itself (Dasgupta 1974, pp. 330-337)

Moreover, even if a difference principle was conceivable for the intergenerational framework, there would be no way to act on the past (Brandstedt 2017, p. 270), that is, the criterion could be applied only from the moment of the “entry in society”[29] onward, while it would be impossible to carry out its prescriptions towards any previous generation. For example, if the least advantaged subjects, after the veil is dropped, were located in the past (in a moment “before” the agreement), there would not be any concrete way to fully realize the difference principle’s (intergenerational) prescriptions[30]. As for this point, Rawls is extremely clear: “there is no way for later generations to help the situation of the least fortunate earlier generation” since “it is a natural fact that generations are spread out in time and actual economic benefits flow only in one direction” (Rawls 1999, p. 254).

For the just mentioned reasons, in Rawls’s opinion the difference principle is not a suitable tool to deal with redistribution of resources across time: “thus the difference

28 In terms of primary social goods, that "are things which it is supposed a rational man wants whatever else he wants" (Rawls 1999 p. 79). See also (Rawls 1999, pp. 78-81)

29 As highlighted many times by Rawls, the original agreement is hypothetical, therefore the words "moment", "before", "after" and so on and so forth have to be taken with the right caution and interpreted coherently with the context.

30 This way of reasoning is in line with the idea that within an intergenerational context “ought implies can” (Partridge 2017) and the worst-off should be accessible (Gaspart et al. 2007, p. 203), no matter the generation they belong .

principle does not hold for the question of justice between generations and the problem [...] must be treated in some other manner" (Rawls 1999, p. 254). Rawls therefore proposes the just saving principle as the normative solution to guarantee intergenerational redistributive fairness. In short, "the difference principle holds within generations" while "the principle of just saving holds between generations" (Rawls 2001, p. 159).

"The just savings principle applies to what a society is to save as a matter of justice" (Rawls 1999, p. 255) and "in following a just savings principle, each generation makes a contribution to those coming later and receives from its predecessors" (Rawls 1999, p. 254). However Rawls does not enter much in detail in describing the particular features of the mentioned principle (like for example providing a specific saving rate or a scheme of rates). Instead he limits himself to sketch some general ethical restriction which the contractual parties should take in account in defining the saving path (Rawls 1999, pp. 255-6).

Nevertheless here it is not of particular interest to linger on those, although reasonable, ethical balances. Instead it is relevant to understand which is the positive reasoning offered by Rawls to substitute the inadequate difference principle with the just saving principle. He essentially proceeds in two parallel steps:

- first of all Rawls restates the intergenerational redistrivutive problem. The parties behind the veil of ignorance are aware of the natural flow of economic benefits (which is a general and unalterable circumstance), therefore the new issue becomes to understand how the generation involved in the agreement can fairly treat not all the possible generations of the history but only the subsequent ones;
- second, to induce the subjects involved in the agreement to think not only as contemporaries but to take into consideration also the future generations, Rawls amends his own theory and adds an intergenerational motivational interest assuming that "the parties represent family lines", that is, they "care at least about their more immediate descendants" (Rawls 1999, p. 255 and Brandstedt 2017, p. 276).

Those are the further specification introduced by John Rawls in order to deal with the peculiarities concerning the distribution of resources between generations within his social contract theory. Thus, reminding how the subjects behind the veil of ignorance are unaware of the historical period they belong and adding to this premise a carefulness for the closer

future generations (the "family line" assumption), the parties involved in the agreement naturally derive the just saving principle[31].

One important feature of the just saving principle is then its duration. In fact, the principle is not required to be applied forever, but the resources have to be moved towards future generations only until the specific task which was designed for is accomplished. In particular "once just institutions are firmly established and all the basic liberties effectively realized, the net accumulation asked [by the just saving principle] falls to zero" (Rawls 1999, p. 255). Thus, "the just savings principle can be regarded as an understanding between generations to carry their fair share of the burden of realizing and preserving a just society" and "the end of the savings process is set up in advance" (Rawls 1999, p. 257)[32].

In this perspective, the Rawlsian intergenerational theory shows to be essentially structured in two distinct stages: a temporary phase of accumulation lead by the just saving principle; a steady state where it is not required to apply any particular intergenerational (redistributive) principle (Gaspart et al. 2007, pp. 193-197, Gosseries 2008, pp. 18-19 and Gosseries 2016, pp. 79-85).

While one fringe of the secondary literature almost uncritically accepted Rawls's conclusions on the just saving principle (Arrow 1973, Dasgupta 1974 and Solow 1974), the majority of the authors showed some perplexities about the approach adopted by Rawls to deal with redistribution of resources between generations.

Thus the critical literature claimed that Rawls substantially failed to apply the veil of ignorance to the intergenerational context (Tremmel 2009, pp. 149-154). With his intergenerational framework Rawls was considered to reach "a modest conclusion" (Heyd

31 Rawls concludes his intergenerational theory adding a really important elucidation concerning the just saving principle, outlined as a formula to represent the duty to sustain just institutions across time. In particular he specifies how "the difference principle includes the savings principle as a constraint" (Rawls 1999, p. 258). That means that before applying the difference principle it is necessary to fulfil the requirements of the just saving principle. Said otherwise "the just savings principle demands that we leave enough capital and resources for future generations while making transfers to our contemporary poor (as required by the difference principle)" (Heyd 2009,p. 171).

32 This structure implies that the just saving principle does not pay any direct attention to the worst-off (like the difference principle does) and more in general it is not concerned with the pure redistribution of resources between generation. Instead, its main aim goal seems to be exclusively to secure the conditions for the realization of just institutions and of a just society (Attas 2009, p. 211, Heyd 2009, p. 187, Gabor 2013 p. 305, Gosseries 2016, p. 80, Paden 1997, pp. 28-29 and p. 38, Wall 2003, p. 93).

2009, p. 187) since he does not provide any particularly elaborated intergenerational distributive theory (Gosseries 2016, p. 87), while for most of the authors Rawls's approach to justice between generation appeared to be limited and unsatisfactory in its deductions (Mathis 2009, Paden 1997 and Partridge 2017).

The general disappointment is then ascribable to different specific critiques. For example, Gardiner (2009, p. 81) claims that the just saving principle does not treat really fairly all the generations because more concern is paid to the future generations. Indeed the accumulation phases might violate the maximin prescriptions (Gosseries 2016, p. 79) so that a very high price might be paid by the first generations (Agius 2006, p. 324). However this is an implicit utilitarian conclusion that Rawls tries to avoid throughout all his contractarian theory of justice (Rawls 1999).

Again, the motivational altruistic assumption is considered an ad hoc construct (Wall 2003, p. 81) that reflects a conception of the good (English 1977, p. 93) and that undermines Rawls's whole theory since it generates tensions between the Rawlsian intragenerational system and his theory valid between generations (English 1977 and Wall 2003): "the postulate of altruistic interest within the original position therefore compromises the whole systematic derivation from contract theory" (Mathis 2009, p. 54).

Furthermore the artificial trick of the family's chains substantially eludes the real intergenerational problem since in some authors' opinions it is not possible to derive an adequate concern for the whole future from the thin interest for the own offspring (Heyd 2009, p. 175 and Mathis 2009, p. 54), because the "concern for the future cannot be understood in individualistic terms" (Norton 1989, p. 151).

Rawls did not keep indifferent to some of those critiques and tried to improve his social contract approach to the allocation of resources between generations. In particular he simplified the framework (Wallack 2006, p. 91), but he did not change the main outcome concerning the just saving principle. Thus, following some hints provided by other philosophers (Rawls 2001, p. 160, footnote 39), Rawls dropped the most controversial hypothesis within his intergenerational system, that is the altruistic intergenerational concern.

In the last version of his intergenerational theory Rawls assumes that the full compliance condition (English 1977, Heyd 2009, p. 179 and Attas 2009 p. 220) is sufficient

to guarantee intergenerational fairness, even if this new framework does not add anything to the substance of the just saving principle.

However, it is worth to remark that this revised way of approaching intergenerational justice is conceived within the ideal theory and it basically applies a Kantian reasoning: now the parties in the original position are intergenerationally disinterested but they "are to ask themselves how much [...] they are prepared to save at each level of wealth as society advances, should all previous generations have followed the same schedule" and "the correct principle, then, is one the members of any generation (and so all generations) would adopt as the principle they would want preceding generations to have followed (Rawls 2001, p.160).

Appendix B: instructions for the experiment in the veil treatment

The instructions for the baseline and the veil treatment are exactly the same for what concerns the intergenerational game. The latter integrate the former only with the agreement phase.

Good morning,

You are about to take part in an experiment on economic decisions. By participating in the experiment you will be able to earn an amount of money that will depend on your decisions and on those of other participants. The decisions you make will remain completely anonymous and no one will be able to associate your choices with your name. During the experiment you will not be allow to communicate in any way with other participants. In case of communication you will be excluded from the experiment without being paid.

We ask you to read carefully the instructions that have been provided to you and which you can consult at any time during the experiment. The instructions will also be read aloud by one of the experimenters. If at the end of the instructions you will have doubts, raise your hand and wait for one of the experimenters to answer to your questions. At the end of the experiment you will be paid privately in cash. To the payment depending on your decision you will earn an extra €3 as show-up fee.

EXPERIMENT

At the beginning of the experiment you will be randomly assigned to a group of 3 participants (you included). The experiment will then be divided into 2 phases and will start from phase 1. However, for clarity, the instructions first show the details of phase 2

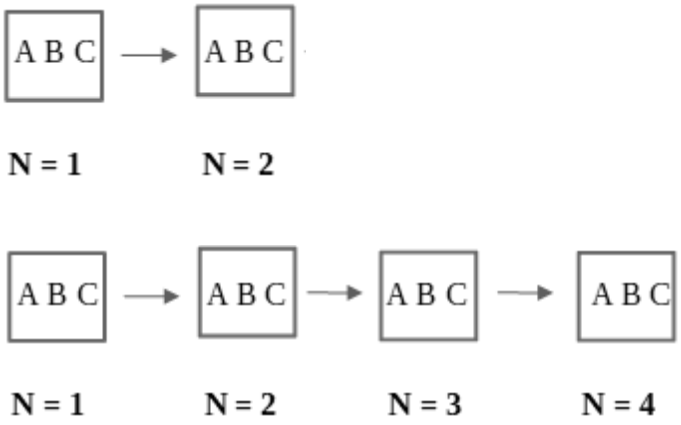
PHASE 2

Only groups that pass phase 1 will have access to phase 2. At the beginning of phase 2, each participant in each group will be assigned an identification letter (A, B or C).

All groups that will access phase 2 will be randomly ordered to constitute various "chains". Within each chain each group of three participants will represent a generation. The chains will have variable length. The shorter chain will be long 1, and therefore will include only one group, but there may be chains of two, three, four, and more groups.

Each group in each chain will be then assigned a number (N) corresponding to the position of the group within the chain. All chains will start with a group in position N = 1, which will represent the first generation of the chain. The second generation, if there will be, will be assigned the number N = 2, the third the number N = 3 and so on for all the groups which compose the chain. Even if you will know the position of your group (N) in your chain, you will not be notified of the total length of the chain, that is you will not know the number of generations that will be there after yours. You will only be certain that within your chain there are other N-1 groups before your group. For example, if your group is assigned to the position N = 3, you know that in your chain there are other two groups in the previous positions, but you do not know how many groups there will be in the subsequent positions.

Below is an example of chains of length 2 and 4



PHASE 2 - DECISION

Each group that will take a decision in phase 2 will face the same type of choice. Participants of the group assigned to the first generation (N = 1) will be the first to make a decision within each chain.

At the time of the decision, the group will have €21 available on a common account. Each participant in the group will be asked to decide how many euros s/he wants to withdraw from the common account by inserting an integer between €0 and €7. The sum that you decide to withdraw on this occasion will constitute your final payment, to which we will add the €3 of the show-up fee. In addition, when you will have to decide how much to withdraw from the account you will not know how much the other members of your group have claimed for themselves.

When all the participants will have made their choices, depending on the amount of euros left on the common account two distinct scenarios might happen:

- If the total amount left on the common account by the group (N) will be at least €6, the chain will continue and the next group (N + 1) will enter the decision phase. In this case the sum left by the group (N) will be completely integrated and on the common account of the next group (N + 1) there will be again available €21.

- If the total amount left by the group (N) on the common account will be less than €6, the chain will be interrupted and the common account will be emptied of any remaining euro. In this case, all participants of any subsequent groups (N + 1, N + 2, N + 3 etc.) will not be able to take any decision and their will be paid only the participation fee of €3.

If your group came at the decision phase, after the choice of how much to withdraw from the common account you will also be asked to make a prediction on the behaviour of the other two participants of your group. You will have to indicate the forecast by entering an integer between €0 and €7 for each of the other 2 participants in your group. The participant of the group who will provide the best forecast will get a bonus of €2 that will be added to their final payment. If two (or three) players provide equally accurate predictions, the bonus will be awarded to both (or all three) participants. The best forecast will be defined according to the following rule (imagining that you are the player A):

SCORE_A= (distance between forecast_on_player_B and decision_of_player_B) + (distance between forecast_on_player_C and decision_player_C)

Thus the SCORE can vary between a minimum of 0 and a maximum of 14. The bonus will be won by the participant(s) whose SCORE will be smaller (or equal) than that of the other two members of the group.

Phase 2 of the experiment will end when all the chains are either exhausted or interrupted.

PHASE 1 - AGREEMENT

In phase 1, along with the other two players in your group formed at the beginning of the experiment, you will have to vote to decide which rule to adopt on the management of the common account in case your group comes to the decision moment in phase 2. However, during phase 1 you will have to vote not knowing which generation (N position in the chain) your group belongs. This information will be provided to you only at the beginning of phase 2. Therefore in phase 1 your group will have to agree on a resource management rule before knowing which N position the group belongs within a chain.

Each subject in your group will have to vote for the rule you prefer, choosing between the following two:

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 more than €5, even if that means interrupting the chain

The resource management rule of the common account must be approved unanimously, that is the agreement will be reached only if all the subjects belonging to the same group have voted for the same rule. In phase 1 you will have 6 rounds to reach unanimity. Groups that

do not reach unanimity within the 6 rounds will not be able to access phase 2 and they will be paid only the show-up fee..

In phase 2 you will then decide whether to apply the rule chosen in phase 1, choosing to withdraw an amount compatible with this rule, or withdraw another sum.

SYNTHESIS

PHASE 1 - In phase 1 you will have to unanimously vote for the rule concerning the management of the common account in phase 2 without knowing which generation (position) of the chain your group belongs. The agreement on a rule is an essential prerequisite to access phase 2.

PHASE 2 - In phase 2 you will know the generation (N) your group belongs in the chain and you will have to decide whether to apply the rule unanimously chosen by the participants of your group in phase 1 or to withdraw a different amount.

Your final payment will therefore depend on the choices made by you and your group during phase 1 and the scenario in which you will be during phase 2 and it will be determined as follows:

- In case your group does not reach unanimity during one of the 6 rounds of phase 1 you will receive only the €3 of the show-up fee

- In the event that your group enters phase 2 but you have not made any choices on how much to withdraw from the common because your chain was interrupted before your group entered could take any decision you will receive only the €3 of the show-up fee

- In case that your group enters phase 2 and you take a decision on how much to withdraw from the common account you will be paid €3 of the show-up fee as + the amount of money

that you have decided to withdraw from the common account + [€2 bonus in case yours is the best forecast]

Before starting with the experiment you will be asked to answer some brief control questions.

References

Agius E. (2006), *Intergenerational justice*, Handbook of intergenerational justice, 53-71, edited by Tremmel J. C., Edward Elgar

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, edited by Axel Gosseries and Lukas H. Meyer, Oxford University Press.

Baggio, M., & Mittone, L. (2018). Grandparents Matter: Perspectives on Intergenerational Altruism and a Pilot Intergenerational Public Good Experiment. Homo Oeconomicus, 1-22.

Bahr, G., & Requate, T. (2007). Intergenerational fairness in a sequential dictator game with social interaction. 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.

Brandstedt E. (2017), The Savings Problem in the Original Position: Assessing and Revising a Model, Canadian Journal of Philosophy, 47:2-3, 269-289.

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.

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, pp 461–471, https://doi.org/10.1007/s10612-015-9293-4

Dasgupta P. (1974), On Some Problems Arising from Professor Rawls' Conception of Distributive Justice, Theory and Decision, 4(3), 325-344.

Degli Antoni G., Faillo M., Francés-Gómez P. & Sacconi, L. (2016), Distributive Justice with Production and the Social Contract: An Experimental Study, Econometica Working Papers, N.60 September.

Dierksmeier C. (2006), John Rawls on the rights of future generations, Handbook of intergenerational justice, 72-85, edited by Tremmel J. C., Edward Elgar.

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. (2008), Compliance by Believing: An Experimental Exploration on Social Norms and Impartial Agreements, available at SSRN 1151245.

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.

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, edited by Axel Gosseries and Lukas H. 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.

Gaspart F., & Gosseries A. (2007), Are generational savings unjust?, Politics, Philosophy & Economics, 6(2), 193-217.

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.

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.

Heyd D. (2009), A Value or an Obligation? Rawls on Justice to Future Generations, 167-188, edited by Axel Gosseries and Lukas H. Meyer, Oxford University Press.

Kahneman D., Knetsch J. L. & Thaler, R. H. (1986), Fairness and the assumptions of economics, Journal of business, S285-S300.

Lamont J. and Favor C., "Distributive Justice", online resource, The Stanford Encyclopedia of Philosophy (Winter 2016 Edition), Edward N. Zalta (ed.), URL =

<https://plato.stanford.edu/archives/win2016/entries/justice-distributive/>, entered August 2017.

List J. A. (2007), On the interpretation of giving in dictator games, Journal of Political economy, 115(3), 482-493.

Mathis K. (2009), Future Generations in John Rawls' Theory of Justice, Archiv für Rechts- und Sozialphilosphie, 95(1), 49-61.

Meyer L., "Intergenerational Justice", online resource, The Stanford Encyclopedia of Philosophy (Summer 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2016/entries/justice-intergenerational/>, entered August 2017.

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.

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. 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.

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.

Partridge E. (2017), Beyond "just savings", online resource, Environmental ethich and public policy, The online gadley, http://gadfly.igc.org/Unpublished/70/bjs.htm, entered July 2017.

Rawls, J. (1963). The sense of justice. The Philosophical Review, 281-305 Vol. 72, No. 3.

Rawls J. (1999), A Theory of Justice, Revised Edition, Harvard Uiversity Press.

Rawls J. (2001), Justice as fairness: A restatement, Harvard University Press.

Rawls J. (2001a), The Law of Peoples: with, the Idea of Public Reason Revisited, Harvard University Press

Sacconi L. & Faillo, M. (2005), Conformity and Reciprocity in the 'Exclusion Game': An Experimental Investigation, University of Trento Economics working paper, (12).

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.

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.

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.

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.

Wolf, S., & Dron, C. (2015). Intergenerational sharing of non-renewable resources: An experimental study using Rawls's Veil of Ignorance (No. 01-2015). The Constitutional Economics Network Working Papers.

Wolf, S., & Wagner, K. (2016). If Future Generations Had a Say: Experimental Evidence on Resource Sharing with Veto Power of a Future Generation.