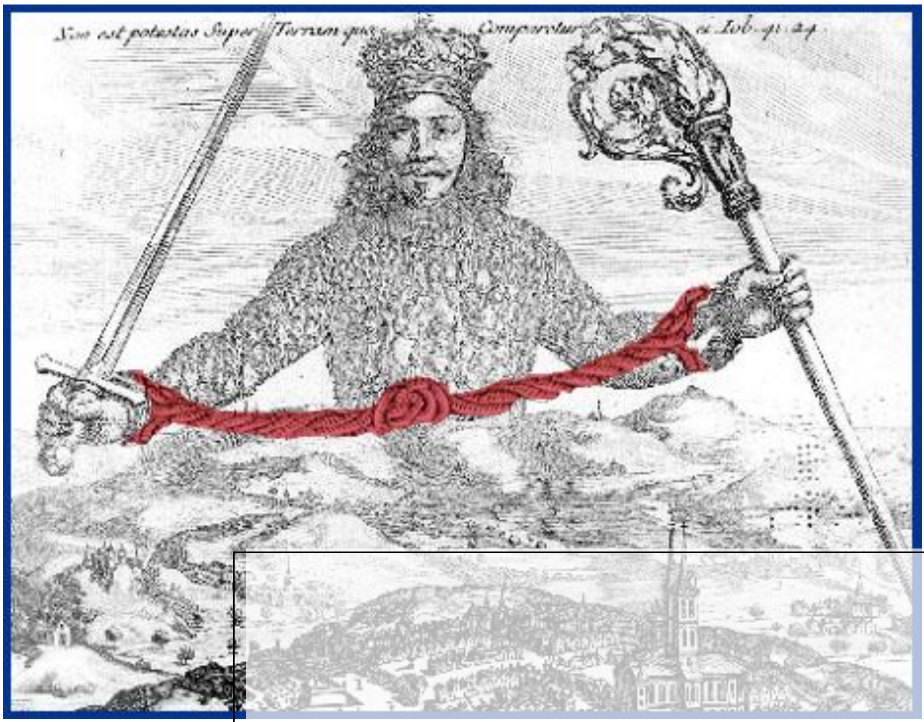

Constitutional
Economics
Network

Working Paper
Series
ISSN No. 2193-7214

CEN Paper
No. 02-2010

An Intergenerational Social Contract for Common Resource Usage: A Reality-Check for Harsanyi and Rawls

Stephan Wolf

Department of Economics and Behavioral Sciences, University of Freiburg, Germany. E-Mail: stephan.wolf@vwl.uni-freiburg.de

July 15, 2010

University of Freiburg
Institute for Economic Research
Department of Economic Policy and Constitutional Economic Theory
Platz der Alten Synagoge / KG II D-79085 Freiburg
www.wipo.uni-freiburg.de

Department of Economics and Behavioral Sciences, University of Freiburg, Germany

Constitutional Economic Network – Working Paper Series

*An Intergenerational Social Contract for Common Resource Usage: A Reality-Check for Harsanyi and Rawls**

Stephan Wolf

Abstract

This paper investigates how one can derive 'fair' allocation shares for renewable and non-renewable resources from a Rawlsian standpoint. Since there are competing interests over limited resources both *within* and *between* generations, it is argued that the respective trade-offs call for a more complete view of the conflict, taking both problems and their interrelation into account. The welfare economic solution of inter- and intra-generational 'sum of utilities' maximization is rejected since it fails to prove that such "optimum" would be chosen by veiled stakeholders in a Rawlsian original position. The individual utility maximizing agent behind the veil is, deprived of knowledge in which generation she will be born and which income group she will part of inside a generation, confronted with a general trade-off: more resources to one generation may improve the lot also for low income individuals, but decrease utility for positions in other generations per se. While the risk neutral agent in Harsanyi's tradition is indifferent between solutions yielding all the same average utility from resource endowment, a realistic degree of risk aversion both concerning the intra- and intergenerational position shape the distributional choice necessarily towards more egalitarian solutions. A crucial factor determining one generation's share of the resource pie for a non-renewable resp. the utilization rate for a renewable resource is discounting. A discount rate of zero, as requested by Rawls, necessarily leads to strictly egalitarian intergenerational regimes independent of inter-generational risk aversion, while the distribution *within* a generation may be, depending on the intra-generational rate, rather unequal. The paper concludes with the observation that unequal inter-generational distributions *among* generations can only be justified given sufficient compensation for resource loss by building up a (public) capital stock.

* The paper was presented at the 2010 conference of the International Society for Ecological Economics, Oldenburg, Germany.

Contents

1 Introduction

Depleting fish-stocks, air atmosphere laden with historically accumulated greenhouse gases, and Damokles' sword in the form of "peak everything" immediately threatening the global economy's vital line–the present global of global common pool resource consumption draws a gloomy picture. How long the mineral and fossil deposits in the Earth's crust will last no-one knows, but the supply of non-dissipated sources of these raw materials is undoubtedly finite. Notwithstanding, renewable commons have been overexploited, often far beyond natural reproduction rates. The use of many common pool resources (CPRs) was an indispensable necessity for creating material riches. However, firstly this production of wealth came with a very unequal distribution of the benefits, and secondly is showing unbearable drawbacks which mainly affect the global poor and future generations. A pressing question therefore is: How can a just distribution of CPRs (and the benefits of their consumption) be reached?

The debate on what is a just distribution is one of the most controversial issues in society. The mere observation that the current situation concerning environmental degradation and resource exhaustion seems unjust does not automatically imply what a just distributional scheme would be. Mainstream economics fails to answer this question. Either mainstream thinkers openly reject any explicit valuation of allocative alternatives and point to politics or political philosophy, or they resort to establishing "efficiency". This in the end is nothing but a hidden normative standard. If welfare economics does not have a satisfactory answer to offer, where could an ecological economist look for useful normative standards for evaluating different allocative regimes, e.g. for CPRs?

One possibility is following the approach of endogenizing distributional norms in economic theory. How this could be done is presented in this paper. This form of deriving normative standards is known as *Constitutional Economics*, but can also be referred to as *Social Contract Theory* (SCT), a well-known strand of political philosophy.

This school of thought goes at back to Thomas Hobbes' "Leviathan" (1651) and Jean Jacque Rousseau's "Contract Social" (1762/2008). More modern representatives of SCT are for example James M. Buchanan (1975), and most famously John Rawls (1971). Both argue, taking up the arguments of their famous forerunners, that societies simply form because of the economic calculus of individuals. The norms ruling these societies are derived from the preferences of those forming a state via consenting to a social contract.

Why may this approach be attractive at all? Especially in the case of Rawls' "A Theory of Justice", the–admittedly hypothetical–construction of (distributional) norms achieved exhibit the quality of categorical imperatives. For skeptics concerning the success of consequentialist normative theories like utilitarianism, on which modern mainstream economics builds up, this might sound quite attractive. Indeed it will be argued in the following paper that there are good reasons why an endogenous derivation of distributional norms from a (hypothetical) social contract approach exhibits desirable properties. Using Rawls' construction of a 'veil of ignorance', I will show that distributional norms can be generated which do not simply reflect the interests of those who are currently in power and benefiting from uneven resource consumption, but which rather request the parties consenting to a social contract to take the position of the poor and the future generations into account. In the end, several possible models are derived which allow the generation of distributional norms, explicitly for the case of renewable and exhaustible CPRs, both among and within different generations. However, this offers a multitude of possible distributional norms. The solution space depends firstly on the formalization of the veil, and secondly on the assumptions made concerning the preferences of the contracting parties (mainly their degree of risk aversion, see Breyer/Kolmar 2005: 41). The good news is: these assumptions can be tested for in experimental settings. In the end, some experiments are suggested which could be run to derive moral principles at least for specific situations of intra- and intergenerational resource distribution.[1]

The remaining paper is organized as follows: in section 2, the standard welfare economic approach to distributional problems is presented and it is argued why it is of little use in reality. Section 3 presents Constitutional Economics and the related concept of SCT as a more appealing alternative. Then in section 4, first the general idea of Rawls' approach is presented. It is shown why Rawls explicitly and openly attacks utilitarianism, and what the general features of his alternative approach are. Special attention is drawn to why Rawls' 'veil'-solution in principle can solve the problem of social contract theory concerning how to incorporate the interests of the unborn, who obviously cannot take part in contractual negotiations. In section 5, the basic arguments of John Harsanyi (cited in Kukathas/Pettit 1990) against the Rawlsian solution are sketched. This is done since Harsanyi explicitly tried to create an endogenous justification of utilitarian thought, which, if correct, would suggest that mainstream welfare economics is the correct normative standard for policy making. Section 6 summarizes the results of a famous experiment conducted in the 1990s to test how realistic Rawls' and Harsanyi's theories are. The result is, briefly, that both ideals are empirically falsified, and these results will play a crucial role for the final part of the paper. Section 7 then presents several formal models which represent the social contracting problem arising from distributional conflicts among and within generations. The novelty is that the standard static modelling of intragenerational conflicts is augmented by a time-dimension, capturing the additional intergenerational conflict. Using different assumptions about individual preferences, a brief comparison of results for a purely Rawlsian, a purely utilitarian, and an empirically more sound social contract are presented. Section 8 concludes with follow-up research questions and policy making.

[1] One drastic conclusion from the overall approach may be that this allows us to pin down the debate over different justice conceptions from a metaphysical level to more 'technical', empirical issues. I myself would not go thus far, especially since also 'empirical' observations are only possible after having accepted certain metaphysical 'believes'. So the problem cannot be ultimately solved, but hopefully at least we gain some direction for further debate.

2 The Welfare Economic 'Solution' to Distributional Justice–and Why It Does Not Solve the Problem

2.1 The Necessarily 'Ad Hoc' Choice of a Social Welfare Function

Welfare economics represents the normative branch of neoclassical economics. With the mainstream economic problem being the 'optimal allocation of scarce resource under constraints', neoclassical economics is all about optimality in the sense of efficiency. More precisely, it is about Pareto-efficiency, which means that an economically efficient solution is reached once all mutually beneficial exchange opportunities have been exhausted. This on the one hand explains why markets (under ideal conditions) always yield Pareto-efficient results: markets are the arena of bilateral exchange, and bilateral transactions will take place as long there are still mutually beneficial exchange potentials. The self-interest then drives rational agents towards a Pareto-optimum–a neoclassical version of Smith's welfare creating invisible hand.
Unfortunately, depending on initial endowments, different Pareto-optimums will result. Hence Pareto-efficiency alone is insufficient to choose between different allocations, each producing one specific distributive result. An additional constraint is hence: a justice criterion. The one and only Pareto-efficient allocation also satisfying this additional condition then has to be chosen as the 'optimum optimorum'(Cullis/Jones 2009)
This immediately raises a follow-up question: where does this criterion come from? Welfare economics itself cannot provide the solution, thus one has to resort to external sources. In "real policy making" this criterion is supposed to be provided by politics. For welfare economists, the remaining task then is to suggest economic policies which make society meet its normative aim in the most efficient way. In technical terms: politics provides a "Social Welfare Function" (SWF), and economists maximize its value by optimal policies. In more theoretical contexts, welfare economists take SWFs as defined by political philosophy (ibid.). Famous examples are the Benthamite SWF and the Rawlsian Maximin-SWF. The problem is: since welfare economics itself does not supply a inherent justice criterion, it also cannot provide consistent criteria to choose between different exogenous possibilities. Whatever choice is made, it necessarily is 'ad hoc'.

2.2 Why Welfare Economists Use a Benthamite SWF as a 'Default'

One may wonder: don't standard models always use one specific SWF which is the sum of individual utilities? Indeed, the 'default' case in welfare economics is to maximize a so-called Benthamite SWF, i.e. the unweighted sums of utilities. The reasons are two-fold (Weimann 2009): First of all, under standard assumptions, a free market system will *decentrally* lead to the maximization of the sum of utilities of all exchange partners. Therefore, using a Benthamite SWF is convenient since it gives a strong argument for markets. Secondly, whatever the distributional characteristics of an allocation, if the sum of utilities (approximated by maximizing aggregate monetary income) at its highest, the 'economic pie' is as large as possible. Consequently there is as much as possible available for after-market redistribution, organized to 'heal' the most extreme inequalities arising from free market allocations. Analytically, allocation and distribution are separated, allowing for a maximum output constrained by an exogenous justice standard.

2.3 Maximizing a Benthamite SWF May Not Be Attractive in Reality

One obvious problem is that *analytically* one might separate allocation and distribution, in reality both result simultaneously from market and state activity. 'Okun's leaky bucket' (ibid.) teaches us: any form of redistribution goes along with a decrease of overall output due to disincentive effects. But even here standard welfare economists find a 'way out': Instead of real compensation via actual transfers, they just suggest those policies which maximize aggregate income and allow for *potential* compensation. Actual losers from policies could be (over-)compensated by policy winners such that everybody is better off (Cullis/Jones 2009). Obviously this is no real solution. Hypothetical compensation does not necessarily lead to *real* transfers. Even if first promised to the worse off, after the welfare maximizing allocation is realized, without enforcement of transfers compensation remains very unlikely. If it is enforced, it is no longer hypothetical and free from disincentive effects.

3 Constitutional Economics and Social Contract Theory: Dissolving the Dilemma

Is there a way to avoid the entire dilemma of the indeterminacy of the 'just Pareto-optimum' and the irrevocability of redistributional trade-offs? According to constitutional economists and social contractarians, there is a true solution. The key is to avoid the application of *exogenous* value standards, but instead to derive distribtutional norms *endogenously* from the preferences of the individuals within a society (Neumï¿$\frac{1}{2}$rker 1998). Of course, assuming this approach is plausible at all, then the question still remains how to find out people's preferences. An exhaustive and infallible way to determine everyone's preferences in totality may not be at hand, but it is possible to make reasonable, empirically sound assumptions on individual preferences.

3.1 The Justification of a Social Contract from a Positive Economic View

It is nevertheless important to point out that this approach, which aims at deriving distributional norms with *economic* tools does not make political philosophy obsolete. On the contrary, endogenizing the choice of distributional norms rather calls for a deep cooperation with political philosophers, since this procedure is not a panacea either. The natural, intellectual brace holding together 'economics of justice' and political ideology is SCT, which in economics is known as 'Constitutional Economics' (CE). SCT is the most 'economical' approach towards explaining the existence and legitimation of political institutions (Rawls 1971). The fundamental idea goes back to Hobbes' Leviathan (1651). Hobbes famously argued that it is in the rational interest of people, who are otherwise caught in an undesirable anarchic state of nature which only allows "the life of man [to be] solitary, poor, nasty, brutish, and short" (ibid.), to surrender all power to a central authority. The emergence of the state and the derived scheme of provision of public goods, above all safety and freedom from plunder, hence gets a purely economic explanation. In a similar way, Rousseau argued for a 'contract social', where individuals join forces for mutual improvement of life. Rousseau drew, opposed to Hobbes an overly optimistic image of the benefits of a society based on a social contract, while Hobbes' Leviathan must look like a horrible construction, only acceptable since it keeps people from killing each other. More modern thinkers rather argue for a more moderate perception of a social contract establishing joint government. James M. Buchanan (1975) highlights the necessity of the state being a 'protective agency' in a Hobbesian sense, but necessarily limited in its extend to avoid a Leviathan government acting against the interests of those having established the very same government. Beyond that, Buchanan also sees potential for a 'productive state' which may contain a redistributional branch, as long as people consent to such on the constitutional, i.e. social contract level.

3.2 One Special Social Contract: the Rawlsian Solution

A more radical version of SCT is presented by Rawls in his famous work "A Theory of Justice" (1973). Not only does he argue that people would consent to some form of redistribution within a generation, but also finds a very specific solution for this. For Rawls, the state is necessary to maintain public order, but it only is desirable if it fulfils certain very clear moral standards. While Buchanan (1975) gives a more pragmatic and positive explanation why a social contract will emerge out of individual interest as such, he does not make reference to any specific political and economic order. Rawls goes as far as to demand that people must be deprived of *all* personal, idiosyncratic information. They are, unlike in Buchanan's theory, informed about the general characteristics of possible future environments, but they know nothing about their own preferences, nor about in which generation they will live and which position they then will hold. Rawls (1971: chapter 40) explicitly follows a Kantian approach to develop moral standards. His argument is that once people are placed behind the veil and thereby deprived of all this information, they are freed from all the related disturbing influences. No longer subject to external influences, people become the autonomous individuals requested by Kant, who then are able to rationally decide on general rules.

4 Rawls' Solution to Distributional Problems: the Just Savings Principle and the Difference Principle

Another interesting trait of Rawls' theory, apart from the veil, is that he explicitly attacks utilitarian thought, which is also the base for today's mainstream economics. One does not necessarily have to follow Rawls' in his–widely criticized (Kukathas/Pettit, 1990)–result that within a generation, rational individuals would adopt the difference principle, also known as the maximin-rule. But at least it is of interest how Rawls argues especially why utilitarian thought and hence also standard welfare economics is to be refuted.

4.1 Welfare Maximization at the Expense of Individual Rights?

The main reason for Rawls to abandon utilitarianism is that it justifies doing great harm to individuals, as long as more benefit is generated for others. Apart from the fact that in reality it is blankly impossible to compare utility levels amongst persons, how can one reasonably justify that someone's basic liberties and rights are infringed for the sake of 'the greater good'? For Rawls no rational individual would ever consent to such a principle which may deprive the very same person of everything she or he has, whatever this 'higher aim' is suppose to be.

4.2 The Intragenerational 'Maximin' Solution

Using the 'veil of ignorance', he tries to show that rational, self-interested (i.e. "mutually disinterested", Rawls 1971) individuals would want to maximize equal political liberties for all, since this is what people in any situation would like to enjoy. Secondly, access to any position in society must be open to everyone. Finally, the distribution of what Rawls calls "primary goods"[2] (1971) must be such that it is to the greatest advantage of the least fortunate group. In total, Rawls develops a comprehensive theory of justice comprising both the political and the economic arena. For the specific distributional question this paper addresses, these aspects are mentioned, but will not be taken into further consideration. The focus indeed lies on how to distribute 'primary goods', to which CPRs have to be counted, since they play a role to create one of the explicit primary goods called "wealth" (ibid.).

4.3 Rawls and Intergenerational Justice: the 'Just Savings Principle'

4.3.1 A First Explanation of Just Savings

Rawls is one of the first political philosopher explicitly addressing intergenerational justice, (Rawls 1971: Chapter 44). He analytically divides distributional conflicts into an intergen-

[2] Rawls remains rather vague on what has to be counted as a primary good but just gives a very general list, above all comprising income and wealth, but his aspect is not very relevant for the remaining discussion in this paper.

erational dimension, and the problem of how to fairly distribute primary goods within one generation. For Rawls the intergenerational problem has to be solved first. The reason for him is that first it must be found out how much of all primary goods are available for each generation. Only then it is fruitful to debate appropriate sharing of one generation's portion of the whole among its current members. Nevertheless, in "A Theory of Justice", Rawls first investigates the problem of intragenerational justice, and the respective solution as rational people would consent to under the veil. His solution is that rational individuals would consent to distributing goods making the least fortunate benefit most. This result is not based on any form of altruism but results from the rational interest of people who have to take into consideration the possibility of themselves ending up in the worst-off position. Since people even then still want to make maximum use of their liberties, this requires that the endowment with the necessary material (and immaterial) goods must be as large as possible. In mathematical terms, Rawls hence derives a 'maximin'-rule or SWF. The overall welfare of society is hence dictated by the utility of the worst off, as derived from the endowment with primary goods of which everyone naturally wants to have as much as possible.

Interestingly, Rawls (1971: chapter 44) rejects this solution for the intergenerational case. The reasons for this are first of all that maximin may not be applicable through time. If one assumes accumulation of wealth over time, which is historically plausible, then one would have to compensate earlier generations who are worse off than later ones by transfers back through time–a logical impossibility. Additionally, an intertemporal maximin-rule may exhibit very undesirable features. If one assumes that mankind progresses over time, then the first generation is the worst off among all. Thus, they have to be made as well off as possible according to Maximin, but this implies that the whole capital stock which is assumed to exist in the beginning will be consumed away, since no-one can request the worst-off to forgo consumption opportunities. Consequently, capital will vanish over time, making progress impossible.

Convinced that 'maximin' would rule out intertemporal saving, Rawls thought about a way out of this dilemma. He came up with two solutions. The first one was described in "A Theory of Justice". It has to be assumed that those behind the veil know that they all belong to the same generation, but also that they represent 'fathers of family lines' (ibid.). Hence we can also assume that each head of family exhibits some degree of altruism and cares for his descendents, and therefore would at least save some fraction of overall capital.

4.3.2 Just Savings as a Reciprocal Decision: Rawls' Later Justification

Rawls was heavily criticized for this first idea, since he ad hoc had to assume altruism for the intergenerational case, whereas he explicitly demanded people to be self-interested in the intragenerational case. Here, many critics claim, a logical inconsistency, or at least argumentative weakness in Rawls theory. Consequently, he reformulated his "just savings principle" in "Justice as Fairness–a Restatement" (Rawls 2000). His new idea was to abandon altruism entirely, but instead to request each generation to respect a savings rate which each generation t would like the previous generation $t-1$ to have adopted, and hence also generation t would follow in the interest of generation $t+1$. Not knowing which generation one will belong to, each rational agent then has an incentive to agree to this "just savings rate". Rawls is not explicit about the actual size of this rate, he just speculates about it being higher or lower depending on the current stock of resources.

In the next section, briefly the critical response to Rawls' ideas from one of his major opponents is briefly presented. This famous critic is John Harsanyi, an economist who tried to defend the utilitarian normative ideal exactly by using Rawls' veil of ignorance.

5 Utilitarianism Strikes Back: Harsanyi's Criticism of Rawls

The lever Harsanyi uses to dismantle Rawls' endogenous explanation of the maximin-rule is his attack on individual decision making under uncertainty as described by Rawls. There is an alternative way to interpret the kind of behavior people exhibit behind the Rawlsian veil: people act as infinitely risk averse behind the veil (Binmore 1989, Roemer 1996). Maximin results then simply due to the fact that people want maximum insurance against the worst case, which is identical to choosing the social state making the worst off as well off as possible–ignoring any potential benefit from becoming any richer than this poorest individual or group. Here, Harsanyi deeply disagrees with Rawls by stating that individuals ignorant of the exact distribution of different social position *nevertheless construct subjective probabilities* about possible events. Lacking any exact information, people then would rationally assume an equal distribution of available positions. This principle is known as *the rule of insufficient reason.* Combined with the standard economic assumption of risk neutrality, the parties behind the veil would consent to maximizing average per capita endowment with primary goods.

For the discussion on the endogenization of SWFs, this is most interesting. Provided that Harsanyi's argumentation is right and Rawls errs, then the use of a standard Benthamite SWF *is* justified, since Harsanyi's endogenous solution coincides with the standard approach in welfare economics: maximizing the exogenously given SWF $W^{Benthamite} = \sum_i u_i(x_i^k)$, where u_i stands for individual $i's$ utility, and x_i^k for the amount of (primary) goods this individual receives in social state k.

6 Who's Right Then? An Empirical Challenge to Both Rawls and Harsanyi

The intermediate result is: Rawls dismisses utilitarianism and presents an alternative distributional theory, while Harsanyi claims that under the Rawlsian veil, rational individuals would exactly opt for the utilitarian principle. The obvious problem is that both cannot be right at the same time. But if the results are not *identical*, then the crucial difference must lie in the different premises both use concerning individual choice given uncertainty in potential outcomes. Both merely speculate on this issue. For a final answer, this situation of course is unsatisfactory. As mentioned in the first section of this paper, if there is dissent about preferences and individual behavior in veil-like situations, then an empirical test may solve the dispute.

Exactly this was the motivation behind the experiment Frohlich and Oppenheimer (1992) conducted in the late 1980s. Both researchers constructed an experimental setting, in which groups had to choose between different distributional schemes. The offers on which the participants had to consent included (a) Rawls' maximin-rule and (b) average income maximization (Harsanyi's solution), as well as (c) average maximization with either an income floor or (d) both floor and ceiling. If people would (nearly) always choose solution (a), Rawls' theory would be supported. If (b) was largely adopted, then Harsanyi's would be right. However, the results Frohlich and Oppenheimer's experiment generated were neither supportive for Rawls nor for Harsanyi. In about 80 percent of the cases, the participants opted for maximizing income combined with an income floor. Harsayni's rule was still adopted in about 15 percent of the cases, which of course is much more significant than one out of 85 groups voting for maximin, but still far behind what Harsanyi's theory would have requested (ibid.).

How can these result be interpreted? The first and most straightforward observation is that neither Rawls' nor Harsanyi's theory performed any close to what each had claimed to be ultimate solution. The experiment was heavily criticized e.g. since only students had participated, and most of the research was done in the U.S., with some cross-checkings in communist Poland. But this criticism fails to understand that *whatever the circumstances of such experiments are*, either Rawls' or Harsanyi's solution should have been stably adopted, since both claim general validity. In terms of risk aversion, one can deduce that people neither opted infinitely risk averse as in accordance with Rawls, nor completely risk neutral as Harsanyi had predicted. The choice of the minimum income of course gave rise to discussion among the participants, since they had some options available, but the overall outcome is quite clear: people wanted some minimum insurance against the worst case (but no total insurance like in Rawls' theory), and then consented to average income maximization–which also is not the same as Harsanyi had suggested, since the maximization was undertaken with a safe minimum income constraint. This experiment provides interesting insights into distributional choices *within* a generation. Since the focus of this paper is, however, how this conflict is related to problems of distributing CPRs *among* generations, appropriate experiments could be conducted, testing the choice of distributional norms both over time and among income groups simultaneously. Before this can be done, the necessary theoretical framework must be developed to derive a meaningful setting.

7 An Intra- and Intergenerational Social Contract

The crucial problem for distributing exhaustible CPRs and utilizing renewable ones is the above described inherent trade-off. Even though one may rather easily come to the conclusion that the current consumption of both renewable and non-renewable CPRs is unfair since it deprives the current poor and the future generations of 'fair development opportunities', it is not clear how these 'fair shares' are to be specified. Is it more pressing and 'more just' to create high economic growth globally such that the current poor are made much better off (via trickle down effects)[3], but which may aggravate the situation for future individuals, both rich and poor? Or does the ecological crisis rather call for maximum protection of the life base of our planet? Although sustainability thinking has been emphasizing more and more that global justice cannot be separated from intergenerational justice, in many concrete decisions this trade-off between the two disadvantaged groups seems unavoidable–at least in the short run. Social contract theory has the potential to derive criteria to solve such conflicts.

Unfortunately, the issue for the intergenerational case is more complex than the intragenerational one. It is generally impossible to redistribute to past generations, a point which Rawls highlights in order to argue why intergenerationally the maximin principle cannot be applied. Additionally, in intertemporal decision making, the issue of discounting future generations' well-being arises. There is a vast literature on whether or not future utilities should be discounted, but no ready-made answer exists. One side argues (see e.g. Rawls 1971) that discounting as such is immoral, since there is no legitimate reason why the wellbeing of those living in the future should be less important than that of the current ones. Others state that the smallest probability of the world coming to an end already makes discounting necessary (LLavador/Roemer/Sylvestre 2009) . Actually people do discount all the time, since future events are uncertain and hence cannot have the same weight as current sure events, and this hence must also hold on the aggregate level (Mueller 1974). I do not want to provide a final judgement here whether or not one should discount future generations' utilities in a social contract setting. Instead I will present several formal models which can be applied in one case or another.

7.1 A General Non-Discounting Model for Diverse Social Welfare Functions

7.1.1 Generalizing SWFs

Cullis and Jones (2009: 8-10) provide a general formal approach to models which are able to capture different SWFs depending on certain parameters. Here, I do not apply the functional form suggested by the two authors, but use a very similar expression presented in Varian (1992). The general features of this function are very similar, especially its ability to express a variety of SWFs depending on the choice of the parameter ρ. The general form of a SWF is then

[3]Even though there is significant doubt about the view that more growth in the global north necessarily finally trickles down to the South; some claim there is rather a conflict, which in the case of CPRs seems obvious (NEF 2010).

expressed as

$$W = \Big(\sum_i (a_i u_i(x_k))^\rho\Big)^{1/\rho}, \tag{1}$$

where $x_{k,i}$ stands for the amount of x individual i receives in social state k, a_i expresses the weight given to individual i's utility. Throughout the paper, it is assumed that $a_i = 1 \; \forall \; i$, since no position is taken to be more important than any other.[4] This expression yields for the two person case with individuals a and b:

$$W = \Big((u_a(x_{k,a}))^\rho + (u_b(x_{k,b}))^\rho\Big)^{1/\rho}. \tag{2}$$

Parameter ρ stands for the degree of risk aversion an individual behind the veil of ignorance exhibits. For $\rho = 1$ (risk neutrality), equations 1 and 2 yield the Benthamite SWF, for $\rho \to -\infty$ (infinite risk aversion) the result is the Rawlsian SWF.

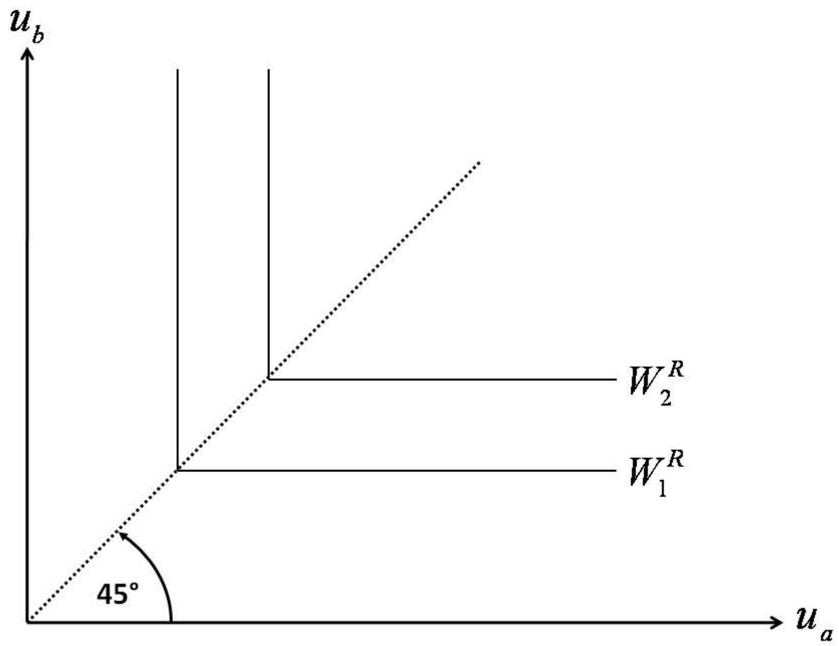

Figure 1: A Rawlsian SWF with no substitution possibilities between individual utilities (based on Cullis/Jones, 2009: 9). The 45ï¿½-line gives the locus of welfare maximums with minimum resource input, the kinked curves represent different social indifference curves (welfare levels). Superscript "R" stand for "Rawlsian". The slope of this curve would change if different weights were given to the two individuals in this two-person example.

Another polar case is complete risk neutrality. This is mathematically equivalent to $\rho = 1$, which yields the well-known standard utilitarian SWF (of Benthamite type, since $a_i = 1 \; \forall \; i$).

[4]This assumption can be relaxed in case individuals have reason to assume that the likelihood to end up in different positions no longer is equal.

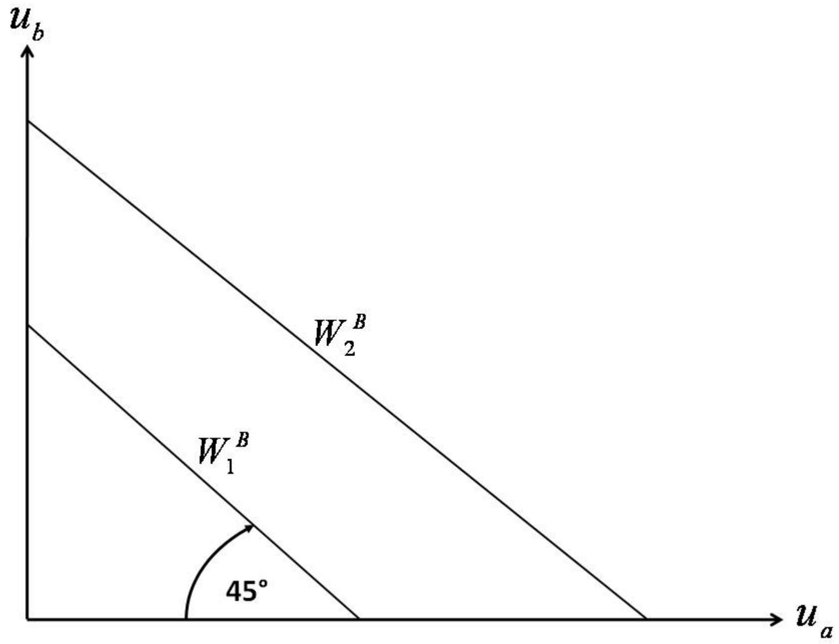

Figure 2: A Benthamite SWF with full substitution possibilities between individual utilities (based on Cullis/Jones, 2009: 9). The 45ï¿½-lines represent different social indifferent curves (welfare levels), the locus of welfare maximums with minimum resource input. Superscript "B" stands for "Benthamite". The slope of these curves would change if different weights were given to the two individuals in this two-person example.

Any degree of risk aversion between neutrality and infinity, which sounds plausible to assume, especially given the results of Frohlich/Oppenheimer (1992), naturally gives a SWF which has properties somewhere between the Benthamite and the Rawlsian type. Substitution as such is not impossible, but not to the degree as in the Benthamite case. Individuals have a stronger inclination to avoid low incomes, but given a certain probability to end up in a better position, this possibility is not completely ignored like in the Rawlsian case. Figure 3 shows how different values for ρ influence the shape of the welfare function for one specific welfare level.

7.1.2 Generalizing the Maximization Problem over Time and Income Groups

The model so far only considers one state of nature, an arbitrary k, taking into account the utility of two individuals. For the problem of inter- and intragenerational distribution (of CPRs), one refinement and one extension are made:

1. Individuals a and b stand for a 'rich' and a 'poor' individual respectively for a representative from each group.
2. The different states of nature k are interpreted as belonging to different generations, so we now use the variable t instead of k, where $t \in T$, and $T \in \mathbb{N}_0$, which stands for the total number of generations. We thus have to sum up over all possible states of nature,

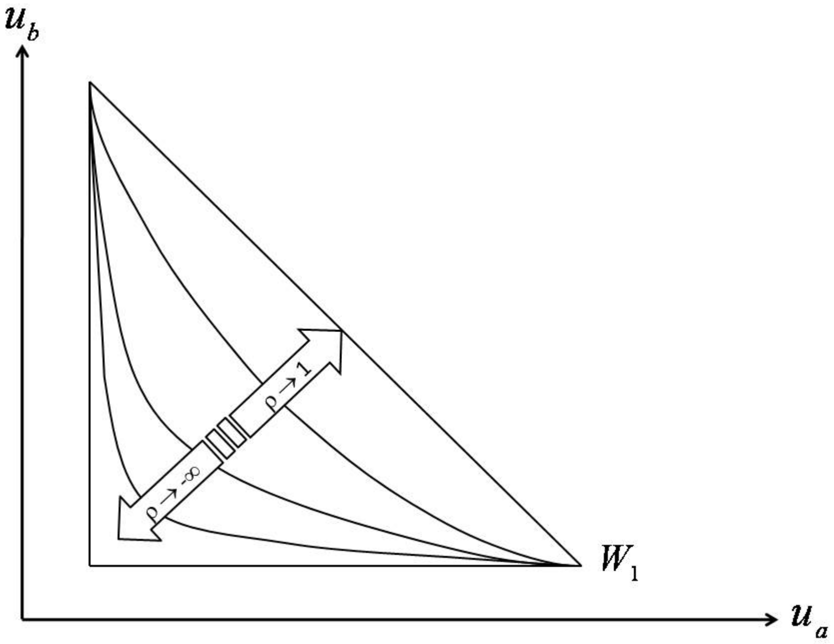

Figure 3: A SWF where people exhibit different degrees of risk aversion, lying between neutrality and infinite aversion (author's illustration).

where belonging to one generation has simply the probability $pr_t = \frac{1}{T}$. A crucial question, which will lead again to the issue of discounting, is whether we take T as finite or let $T \to \infty$.

The overall first model then looks like the following:

$$W(T) = \left(\sum_{t=0}^{T} \left(\left(\sum_{i} (u_i(x_{t,i}))^{\rho} \right)^{1/\rho} \right)^{\phi} \right)^{1/\phi}, \tag{3}$$

or for the case of having simply "rich" and "poor":

$$W(T) = \left(\sum_{t=0}^{T} \left(\left((u_r(x_{t,r}))^{\rho} + (u_p(x_{t,p}))^{\rho} \right)^{1/\rho} \right)^{\phi} \right)^{1/\phi}. \tag{4}$$

The standard, but crucial assumptions about utility functions are $u_i'(x) > 0$ and $u_i''(x) < 0$. In addition to ρ (the degree of risk aversion *within* a generation) we now have $\phi \in (-\infty, 0)$ being the degree of risk aversion *among* generations. An individual behind the veil hence has to take into consideration that she not only might end up being rich or poor within one generation, but also that she might be born at different points of time, represented by t. Becoming the rich one in one specific generation then might be less attractive than being a poor one of another generation if the total income of the former generation is sufficiently larger then of the latter one. When deciding on how to distribute a finite amount of a CPR or how to extract the product of a renewable pool resource, the individual now has to take both the time and the intragenerational dimension into consideration.

7.1.3 A First Result for the General, Non-Discounting Model

This already gives us a first result: Whatever value ϕ takes, the equal probability of ending up in one of the generations dictates that for the exhaustible CPR, the distribution of the total amount $\bar{x}$ must be equal over all generations. Thus, with x_t denoting the amount of our CPR x distributed to each generation t, every generation receives $x_t = \frac{\bar{x}}{T}$. The remaining question then is how much will be distributed to the poor and the rich in each generation. The higher ρ, the more equal the distribution within a generation has to be, while for complete risk neutrality any distribution from giving nothing to the poor to equal distribution (assuming that "poor" always means having less than the "rich") is possible. So as long as there is no additional constraint on intragenerational distribution, the solution may be undetermined.

7.1.4 Changes to the Solution in Case of Redistributional Inefficiencies

One interesting extension of the model would be to assume that the CPR does not directly yield utility, but that it first must be transformed into some consumption good, where the production of this good also requires labor input which can be exerted both by the rich and the poor. "Rich" and "poor" than should be interpreted as having more respectively less possibilities to efficiently transform the CPR into the consumption good, so we could interpret "rich" as 'more talented, technologically more equipped' and "poor" as 'less talented, technologically under-equipped'. This on the one hand may capture the fact that the educational levels in poorer countries are usually lower, and on the other hand the endowment with productive capital is inferior. It therefore would be more efficient to produce the consumption input in the rich region, such that the amount of the consumption good is higher given a certain value of x_t. Consequently, the distribution question arises. If production has taken place in the richer countries, aren't then those having exerted the labor effort the owners of the product? Or should redistribution be organized along the lines of need, which may be even larger in poorer countries? How is the fact taken into account that the CPR production input may be collectively owned?
Whatever the result of this debate, redistribution from the more to the less productive group creates the disincentive effects described in section 2. Consequently, there is a leaky bucket problem. For the individual behind the veil, the relevant trade-off then is how much overall product to let forgo for how much 'insurance against the worst case'. More equal distributions then go along with a lower *total* income, but with a higher *fraction* of the total product received by the poor. The size of ρ hence will determine, depending on the inefficiency parameters and the disutility from work, how much total output an individual will be ready to lose for making the worst-off sufficiently well off. Intragenerationally, we are hence confronted with a true trade-off. This problem might also arise intergenerationally, in case earlier production creates

positive or negative effects on the wellbeing of entire generations.
For infinite risk aversion, the solution would be for sure to guarantee equal incomes, even if this comes at maximum inefficiencies. Risk neutrality would lead to no redistribution at all, since in this case overall production is maximized, hence disincentives do not arise. So even without rigorous mathematical derivation, this first model gives us intuitive reasons for one or another distribution of exhaustible CPRs. One problem arises in the case of such resources: For $T \to \infty$, the model obviously yields that $x_t \to 0$. The whole exercise then gives a trivial, non-satisfying result that the CPR will not be used at all. Thus no distributional problem arises, neither between nor within generations. This is of course nothing new and has been long discussed in resource economics. Two solutions may lead out of this problem:

1. Discounting the utility of future generations leads to $x_t \to 0$ for $T \to -\infty$, with $x_t > 0$ and $x_t > x_{t+1} \quad \forall\, t$.

2. If the CPR x helps building up a capital stock and does not directly yield utility, then complete depletion of x over time may be justified, as long as the production input x is at least partially substitutable (x is hence not an essential resource).

The solution actually also changes in case the CPR is renewable: for the case T is finite, the resource stock in the end will be depleted entirely, which drives the reproduction of x to zero in the last period. For $T \to \infty$, the non-discounted sum of this model suggests utilizing the resource at the maximum sustainable yield, such that the stock is maintained at the necessary level to maximize production. The system hence is sustainable in the sense of running *ad infinitum*, with the remaining question of how to distribute the product within a generation.
In the next part of this section, the problem of how to interpret the 'just savings rate' is investigated.

7.2 Two Rawlsian Models for Intergenerational Distribution

7.2.1 'Rawls I'

Even though Rawls explicitly rules out discounting between generations, there is nevertheless a possibility to introduce discounting in the intergenerational Rawlsian framework. Why, one may wonder, should a model based on Rawls' earlier, but dismissed works, be created at all? One argument is: allowing some form of discounting (which may be justified from an empirical viewpoint, depending on the experimental setting) yields the interesting result that maximizing an infinite stream of discounted utilities, which is the standard welfare approach, can be explained also in Rawlsian terms! Murrel (1980: 17) shows that as long as individuals exhibit some degree of altruism with respect to their immediate off-spring (the first Rawlsian idea of heads of families behind the veil), those behind the veil know that they are all part of the same generation and face the following maximization problem (where there is only one representative individual per generation, hence no 'rich' and 'poor'):

$$Max\,W = u(x_t) + \beta u(x_{t+1}), \tag{5}$$

with $\beta \in (0; 1)$ being the degree of altruism (Murrell 1980).
Instead of writing $u(c_t)$ as in the original, I have instead used $u(x_t)$. This does not create any problem, since it is assumed that x_t either directly yields utility, or is transformed into a consumption good c, where again marginal utility from consumption is positive and decreasing. The use of x just makes notation consistent.
Murrel concludes that this way any generation t *indirectly* takes into account the utility of all future generations, but *discounted* by the factor β. Starting from time t, the infinite utility maximization problem for one generation hence wriis (ibid.)

$$Max\,W = \sum_{s=0}^{\infty} \beta^s u(x_{t+s}). \tag{6}$$

Even thought there is no discounting factor stemming from time (the value is therefore independent from t), altruism *de facto* works as a discounting factor. The conclusion is that the initial Rawlsian idea ('Model Ralws I') would indeed imply discounting. Vice versa, discounting can, but doesn't have to be derived from Rawls' initial justification of a 'just savings rate' . The exhaustible CPR would hence be consumed at higher rates in the beginning, and at decreasing yearly consumption rates be brought close to extension over time. Also the renewable resource would in the end suffer from complete exhaustion, since maintaining the capital stock has no value at an infinitely far away point of time.

7.2.2 'Rawls II'

Since Rawls (Rawls/Kelly 2001) himself dismissed his own initial approach, I shall end the discussion here. It suffices to see that at least for the intergenerational case, 'Rawls I' would actually imply a discounting solution, despite Rawls' arguments against putting lower weights on later generations.
The more modern version of Rawls' 'just savings principle' is to be expressed in terms of a transfer of resources from one generation to the next. Rawls (ibid.) has basically in mind a stock of depreciating capital which first has to be build up and then can be consumed and/or maintained. The idea is that an intergenerational equilibrium will result where each generation transfers (saves) a certain amount of the capital stock at an amount which the very same generation would have liked to receive from its predecessors. The savings rate hence will reach a constant level over time where each generation just saves as much as necessary to maintain the capital stock at a certain level. It is straightforward to see that this level will be the equilibrium where the marginal utility from saving equals the depreciation rate. The problem is: how can

we assure that a positive amount of the capital stock will be built up to begin with? Since no generation knows when it will be born, but there is a large number of generations, the problem of ending up in the first generation can be neglected, and hence a constant savings rate will result. Then the earlier generations will build up the capital stock (since they consented to this path under the veil), until the maximum level of K is reached where the just savings principle is met.

Formally, this ('model Rawls II') can be expressed as

$$MaxW = \sum_{t=0}^{T} u(x_t - x_t^s + x_{t-1}^s) = \sum_{t=0}^{T} u(x_t) \tag{7}$$

with x_s^t the amount of x saved by generation t for the next period, and x_{t-1}^s the savings received from the last period. According to Rawls', it must be that $x_t^s = x_{t-1}^s = x_{const}^s$ for all generations. For the case of an exhaustible CPR, the problem is not of building up a capital stock, since the stock is there, but rather of how to achieve a rate of maintaining a certain stock level over time. So the level of x being transferred over generations must be constant, or in mathematical terms: $x_t = x_{const} \; \forall \; t$. For $\bar{x} > 0$, this means that the CPR stock of an exhaustible resource never is touched and just maintained constant over time (whatever T, be it finite or infinite). For a renewable CPR, the optimal solution is to transfer the amount of x yielding the maximum sustainable yield. In each case, intragenerational distribution remains to be solved, but is organized according to the respective distributional norm derived behind the veil.

7.3 Summary of the Distributional Results

Table 1 summarizes the above discussed intergenerational distributive results.

Non-Renewable Resource		Renewable Resource	
No Discounting	Discounting	No Discounting	Discounting
$x_t =$ const. $x_t \to 0$ for $T \to \infty$	$x_t \to 0$ for $T \to \infty$ $x_t > x_{t+1}$	$x_t =$ MSY Flow and Stock constant	$x_t \to 0$ for $T \to \infty$ Stock will be depleted over time.

Table 1: Distributional results for renewable and non-renewable resources over time, both with and without discounting (source: author's compilation).

The intra-generational distributional results depend first of all on the size of the trade-off between efficiency and equality-inducing redistribution and the extent of risk-aversion (size of ρ) individuals exhibit behind the veil of ignorance.

Figure 4 graphically presents the dynamic path of the resource stock for (1) a renewable and a (2) non-renewable resource, both in case of (a) no discounting and (b) discounting.

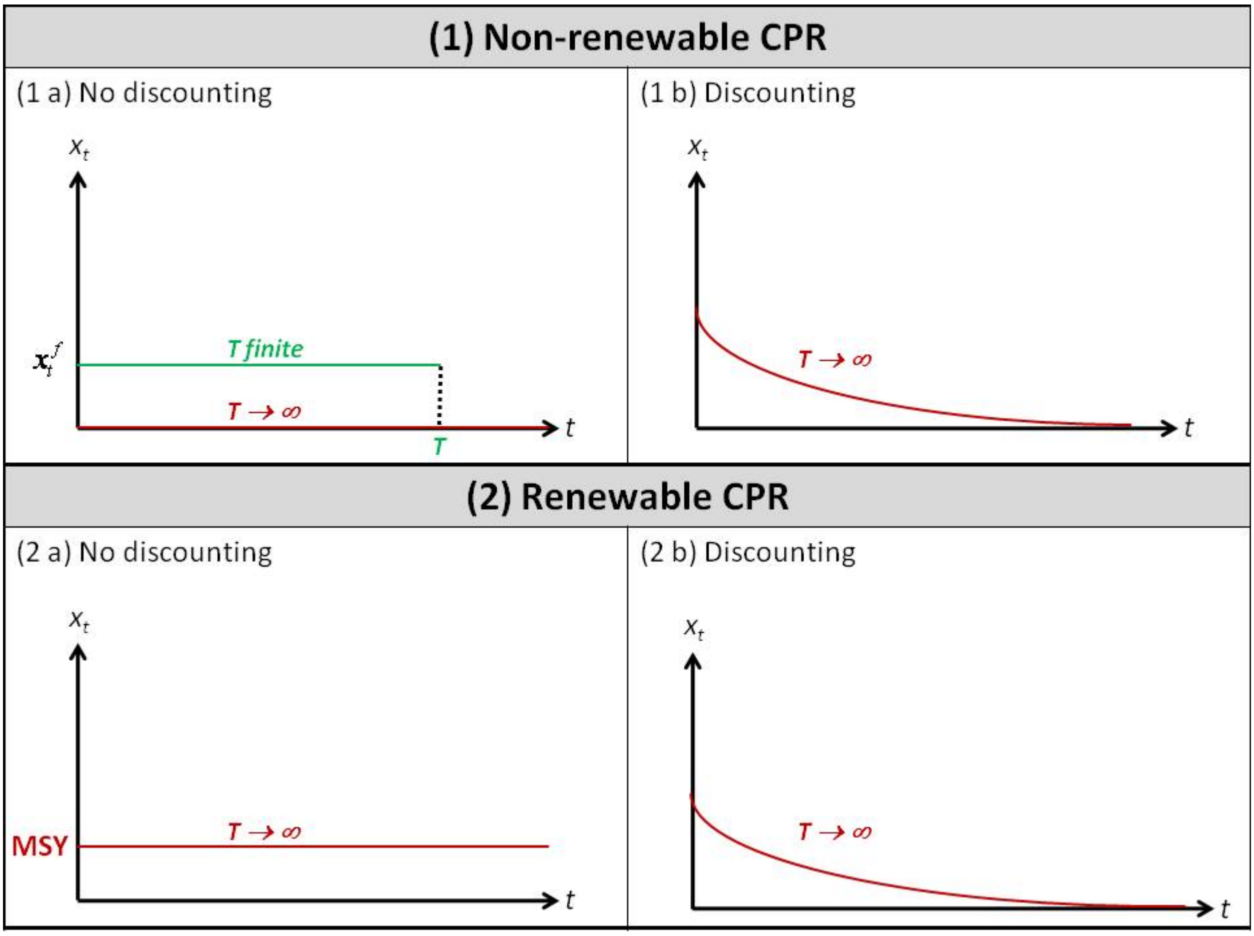

Figure 4: Resource stock paths for (1) a renewable and a (2) non-renewable resource, both in case of (a) no discounting and (b) discounting. Note that path (2b) is flatter, assuming the same initial stock quantity $\bar{x}$, than path (1b) because of x reproducing in case (2). Nevertheless, the stock will be nearly depleted unlike in the non-discounting case since it is rational for discounting individuals to consume not only the flow but also the stock whose value approaches zero for T going to infinity (author's illustration).

8 Conclusion and Outlook

Applying the theoretical results from Harsayni and Rawls to some models based on a 'veil of ignorance' framework, it was shown that in case of exhaustible and renewable CPRs, the distributional results may sometimes display undesirable results, and sometimes suggest highly egalitarian solutions–at least in one dimension.
What remains to be investigated is:

- How will the theoretical solution look if a formal trade-off is integrated into the models concerning intragenerational distribution? How may it depend on the degree(s) of risk aversion?
- What happens to exhaustion paths if the CPRs are used to build up a capital stock?

What size will this stock have, and where could a 'just savings rate' lie?

- Is there a way to empirically test the validity of the different models?

Concerning the first two questions, the task is simply to augment the model by the required constraints and formal relations. About the last issue, I suggest to create experimental settings similar to the one of Frohlich/Oppenheimer, naturally trying to avoid the main mistakes attributed to their investigation. What has to be included in the new setting is an explicit time dimension. Even though the first model yields non-sense results (or at least not attractive ones) for $T \to \infty$, fixing the number of generations to some 5 or 10 generations, with at the same time having rich and poor individuals in each generation, one could indeed try to identify the degree of intertemporal risk aversion. For the case that in the setting the pay-outs (which must be proportional to the respective CPR amount) are handed to the participants immediately after conducting the experiment, one should not expect the parameters ϕ and ρ to differ significantly. Therefore, the pay-offs for different generations should be given to later generations significantly after the experiment. For example, generation t could receive the bonus with a delay of t days or weeks. That way, it should be possible to separate both rates of risk aversion, and to check whether people tend to actually discount utilities, which may suggest tending towards the second model ('Rawls I'). As another research strand, a capital accumulation model like in 'model Rawls II' could be designed, where the explicit task for the participants is to agree on a 'just savings rate', and later to solve the intragenerational distribution problem as suggested by Frohlich/Oppenheimer.

At this stage of discussion, no *direct* policy recommendations are derived. The described approach nevertheless is not useless. First of all, it allows a more precise identification of the distributional conflicts in CPR usage. Secondly, compared to standard welfare economics, the SCT framework provides normatively much more attractive benchmarks for policy making. This holds especially for the intergenerational case, because the veil approach enables to simulate the situation of future people as well, *a fortiori* if backed up by experimental data. Making decisions on behalf of future individuals then still remains a speculative business, but since this is indispensable, a much closer approximation to the interests of later generations is already a step forward.

References

Binmore, Ken (1989). "‘Social Contract I: Harsanyi and Rawls"’. In: *The Economic Journal* (99/395). S. 84-102.

Breyer, Friedrich; **Kolmar**, Martin (2005). *Grundlagen der Wirtschaftspolitik.* Mohr Siebeck, Tï¿½bingen.

Buchanan, James M. (1975). *The Limits of Liberty.* University of Chicago Press.

Cullis, John; **Jones**, Philip (2009). *Public Finance & Public Choice.* Oxford University Press.

Frohlich, Norman; **Oppenheimer**, Joe A. (1992). *Choosing Justice. An Experimental Approach to Ethical Theory.* University of California Press, Berkeley.

Hobbes, Thomas (1651/1997). *Leviathan.* Wissenschaftliche Buchgesellschaft, Darmstadt.

Kukathas, Chandran; **Pettit**, Philip. (1990). *Rawls. A Theory of Justice and its Critics.* Polity Press, Cambridge.

LLavador, Humberto; **Roemer**, John; **Silvestre**, Joaquim (2009). *Intergenerational justice when future worlds are uncertain.* URL: http://www.econ.upf.edu/docs/papers/downloads/1178.pdf [July 1, 2010].

Mueller, Dennis (1974). "Intergenerational Justice and the Social Discount Rate". In: *Theory and Decision.* Ed. 1974: 5.

Murrell, Peter (1980). "The Use of the Maximin Principle as a Criterion for Just Savings." In: *Scandinavian Journal of Economics*, 1980: 2.

Neumï¿½rker, Karl Justus Bernhard (1995). *Finanzverfassung und Staatsgewalt in der Demokratie.* Lang, Frankfurt am Main.

New Economics Foundation (2006). *Growth isn't working.* URL: http://www.neweconomics.org/public growth-isn't-working [June 18, 2010].

Rawls, John (1971). *A Theory of Justice.* Harvard University Press, Cambridge.

Rawls, John; **Kerry**, Erin; (2001). *Justice as Fairness.* Harvard University Press, Cambridge.

Roemer, John E. (1996). *Theories of Distributive Justice.* Harvard University Press, Cambridge.

Rousseau, Jean-Jacques (1762/2008). *Vom Gesellschaftsvertrag.* Reclam, Stuttgart.

Varian, Hal (1992). *Microeconomic Analysis.* Norton & Company, New York.