DEPARTMENT OF ECONOMICS

Working Paper

Intergenerational Justice in the Hobbesian State of
Nature

By

Paola Manzini, Marco Mariotti and Roberto Veneziani

UNIVERSITY OF MASSACHUSETTS AMHERST

Intergenerational Justice in the Hobbesian State of Nature[*]

Paola Manzini[†]   Marco Mariotti[‡]   Roberto Veneziani[§]

This version: July 2010

Abstract

We analyse the issue of justice in the allocation of resources across generations. Our starting point is that if all generations have a claim to natural resources, then each generation should be entitled to exercise veto power on the unpalatable choices of the other generations. We analyse this situation as one of bargaining à la Rubinstein, Safra and Thomson [15], which incorporates a notion of justice as mutual advantage, rather than justice as impartiality, as in the Kantian-Rawlsian tradition. Our framework captures some key aspects of the interaction between isolated agents in a Hobbesian state of nature, in which agents are not placed behind a veil of ignorance, but none of them is sufficiently strong to impose their will against all others (state of war of all against all). We analyse some new social welfare relations emerging from this Hobbesian framework.

JEL: D63 (Equity, Justice, Inequality, and Other Normative Criteria and Measurement); Q01 (Sustainable Development)

Keywords: Intergenerational justice; bargaining; Hobbes; social choice.

1 Introduction

How should resources be allocated across generations? If there is an exhaustible resource that each generation has the power to run down completely, what is the fair proportion that should be saved for future generations?

The interest of each generation clashes with the interest of future and past generations, and the normative question of how to best solve this problem is complex and still open, in the sense that various requirements that are appealing at the normative level are mutually incompatible.[1] One stream of this literature puts at its core an economy in which each generation tackles the problem of how to distribute the available resources between the present cohort of agents and those that will come in the future. The other instead considers the problem of ranking infinite utility streams without investigating how these have come about. In essence this latter approach favours an agnostic position vis a vis the unknown possibilities and avenues that future technological progress might open (or close), focusing on the "purest" problem of intergenerational equity where the issues concerning how the available resources could be invested/preserved are eschewed. One consequence of this is the lack of 'discipline' in the type of infinite utility streams that a planner might have to compare - such general setup makes the problem of impossibilities even worse.[2]

Our contribution lies in this second strand of the literature, from which we depart by dropping the usual requirement of justice as fairness or impartiality, as in the Kantian-Rawlsian tradition, in favour of a notion of justice as mutual advantage (for a discussion of these two perspectives in theories of distributive justice, see Barry [4]), by adopting a *bargaining framework.*

---

[1] If we take fairness across generations as a basic tenet of interegenerational justice, the requirements of Pareto efficiency and anonimity clash with being able to compare any two infinite utility streams (see, for example, Diamond [8], and Arrow [1] for the maximin principles, or more recently Hara, Shinotsuka, Suzumura and Xu [11]. Asheim [2] is a thorough recent review of this literature). Indeed, there are various normatively appealing axioms that are mutually incompatible, in the sense that there is no social welfare function that satisfies all of them. Recently, Asheim, Bossert, Sprumont and Suzumura [3] (who also contain helpful references for the above literature) have promoted a new approach that in a way relaxes the need for completeness: rather than attempting to rank all possible infinite utility streams, they opt for a choice theoretic approach, and what is required is a social welfare *choice* function that can pick the most desirable utility stream among any collection of alternative vectors. This setup is fruitful, as e.g. some of the aforementioned impossibilities are resolved.

[2] See, for example, Ferejohn and Page [9] or Bossert and Suzumura [6].

Our starting point is that if all generations have a claim to natural resources, then each generation should be entitled to *exercise veto power* on the (from its point of view) unpalatable choices of the other generations. Technically, we translate this into adopting as a modeling tool a multiperson version of the bargaining framework à la Rubinstein, Safra and Thomson [15], which we apply to an intergenerational context. As an alternative interpretation, we analyse the Social Welfare Relations (henceforth, SWR) that would emerge from the interaction between individuals in a Hobbesian state of nature. The key features of the Hobbesian state of nature are that, first, individuals are interconnected but they do not cooperate (actually, they are in a state of war of all against all). Indeed, agents are isolated - in Hobbes's own terms, their lives are 'solitary' (*Leviathan*, chapter xiii) - and do not form coalitions. Second, agents are not placed behind a veil of ignorance (unlike in the original position of the contractarian tradition), but none of them is sufficiently strong to impose their will against all others.

We begin by identifying a new SWR that satisfies the requirements above, which we dub 'Hobbesian Social Welfare Relation', and study its properties.

An advantage of using this novel relation as the basis for welfare is that it can handle infinite numbers of agents easily, thus fitting naturally an intergenerationl context. However, as we will see, it comes with some shortcomings. Similarly to the unpalatable behaviour of the standard utilitarian approach to welfare, it is possible that the Hobbesian SWR ranks a profile which privileges a single individual at the disadvantage of many others above a profile where all individuals but the privileged one have high levels of utility. Like for the standard utilitarian approach (e.g. Blackorby, Bossert and Donaldson [5]) the effects of the Hobbesian SWR can be tempered by requiring some form of protection of future generations. Indeed, more generally, our analysis highlights some deep, and perhaps surprising, connections between utilitarianism and the Hobbesian SWRs emerging from bargaining in the state of nature.

However, there are other shortcomings that cannot be accommodated in a similar fashion, and that require us to change the Hobbesian SWR in a more substantial way. This analysis leads us to the formulation of a modified, anonymous Hobbesian social welfare ordering, which can be interpreted as mediating between the two different views of justice as

impartiality and justice as mutual advantage.

We offer no definite conclusions at this stage. The Hobbesian SWR is interesting in its own right as it captures some reasonable normative intuitions, in various morally relevant contexts. Our contribution, however, can also be seen as an exploration of the implications of bargaining approaches to justice. The philosophical foundations and the implications of approaches focusing on justice as mutual advantage have been questioned (see, for example, Barry [4] and Roemer [13]). Critics have also argued out that the approach is ill-suited to deal with intergenerational justice, since removed generations cannot benefit from interaction. Yet the bargaining approach to justice is one of the most influential traditions in political philosophy and its main exponents have consistently applied it to intergenerational problems (see, most notably, Gauthier, [10], chapter IX, section 6). It is therefore of clear theoretical interest to analyse systematically the implications of a bargaining-theoretic approach to intergenerational justice.

The rest of the paper is organised as follows. In section 2, we describe the basic analytical and conceptual framework of our approach to justice as mutual advantage in the intergenerational context. We show how noncooperative bargaining between generations characterises a Hobbesian SWR. In section 3, we then study the properties of the Hobbesian SWR, focusing in particular on intergenerational issues, and the connections between our approach and utilitarianism. We also discuss how the bargaining framework can be modified to accommodate some focal issues in debates on intergenerational justice, such as discounting and sustainability. Some shortcomings of the Hobbesian SWR are also highlighted which lead us to propose, in section 4, a modified Hobbesian social welfare relation which satisfies anonymity and transitivity. Section 5 concludes.

2   Bargaining in the State of Nature

Let N be a set of agents. N can be composed by an infinite number of members. We think of the members of N as 'generations', but the framework applies to other contexts. Let $\mathcal{R}^n$ denote the n-dimensional Euclidean space. Let $\succcurlyeq$ be a (binary) relation over $\mathcal{R}^n$. For any $p, q \in \mathcal{R}^n$, we write $p \succcurlyeq q$ for $(p, q) \in \succcurlyeq$ and $p \not\succcurlyeq q$ for $(p, q) \notin \succcurlyeq$. The asymmetric factor

$\succ$ of $\succcurlyeq$ is defined by $p \succ q$ if and only if $p \succcurlyeq q$ and $p \not\succcurlyeq q$, and the symmetric part $\sim$ of $\succcurlyeq$ is defined by $p \sim q$ if and only if $p \succcurlyeq q$ and $q \succcurlyeq p$. A relation $\succcurlyeq$ on $\mathcal{R}^n$ is said to be: *reflexive* if, for any $p \in \mathcal{R}^n$, $p \succcurlyeq p$; *complete* if, for any $p, q \in \mathcal{R}^n$, $p \neq q$ implies $p \succcurlyeq q$ or $q \succcurlyeq p$; *transitive* if, for any $p, q, r \in \mathcal{R}^n$, $p \succcurlyeq q \succcurlyeq r$ implies $p \succcurlyeq r$. $\succcurlyeq$ is a quasi-ordering if it is reflexive and transitive, while $\succcurlyeq$ is an ordering if it is a complete quasi-ordering. Finally, the vector notation is: $x > y$ (resp. $x >> y$) iff $x_i \geq y_i$ for all i and $x \neq y$ (resp. iff $x_i > y_i$ for all i).

Our objective is to identify a SWR $\succcurlyeq$ on $\mathcal{R}^n$, emerging from a situation of 'social bargaining' over welfare allocations. Here $\mathcal{R}^n$ denotes the set of welfare allocations among agents, or generations, so that if $p \in \mathcal{R}^n$, then p_i represents the welfare level of agent $i \in N$. We imagine a state of nature in which agents are all interconnected, but isolated and cannot form coalitions, as in the Hobbesian state of war of all against all. Consistently with Hobbes's theory, agents are assumed to have equal bargaining strengths and bargaining takes place as in a multiperson version of Rubinstein, Safra and Thomson's [15] interpretation of the Nash Bargaining solution (see also Roemer [13]).

Formally, assume that gains/losses are interpreted additively. Given a proposed profile p, suppose that somebody is willing to pay a cost $e > 0$ in order to *object to* p and propose a different profile q instead. The objection is invalid if there is somebody who is willing to pay the same cost $e > 0$ in order to counterpropose to go back from q to p. The cost $e > 0$ is arbitrary: so we would like to declare a profile p as unobjectionable against q if any objection to p meets a counterobjection of the type above, no matter what $e > 0$ applies. More precisely:

Definition 1 *Allocation p is unobjectionable against q if, for all $e > 0$, whenever there is $i \in N$ such that $q_i - e > p_i$, there is a $j \in N$ such that $p_j - e > q_j$.*

This form of bargaining is an adaptation of two-person games analysed by Rubinstein, Safra and Thomson [15]. Two points are worth noting about the structure of the bargaining process. First, one may question the assumption that agents cannot pool (or transfer) resources in order to make counterproposals. Yet, this is theoretically consistent with the interpretation of the bargaining scenario as a stylised representation of the Hobbesian state of

nature in which we imagine generations to be engaged in. All agents are interconnected but isolated and noncooperative (actually, in war against each other). This conceptual structure of the state of nature underlies the 'individualistic logic' of the bargaining process. However, some libertarian (and therefore strongly individualistic) interpretations of the Lockean state of nature would also be consistent with the logic of the bargaining scenario.

Second, Rubinstein, Safra and Thomson [15] used multiplicative, rather than additive costs. If welfare is level comparable, additive costs are justified. However, a multiplicative version of the same welfare criterion would emerge if only percentage gains/losses could be interpersonally compared.

It is easy to see that the following holds:

Proposition 1 *For any $p, q \in \mathcal{R}^n$, p is unobjectionable against q if and only if*

$$\max_{i \in N}(p_i - q_i) \geq \max_{i \in N}(q_i - p_i).$$

To see this, observe that by Definition 1, p is unobjectionable against q, if for all $e > 0$, for all i in N such that $q_i - p_i > e$, there is j in N such that $p_j - q_j > e$. This can be written as follows: for all $e > 0$, $\max_{i \in N}(q_i - p_i) > e \Rightarrow \max_{i \in N}(p_i - q_i) > e$. The desired result follows from the latter inequality.

3 The Hobbesian Social Welfare Relation

3.1 Some key properties

The previous analysis suggests that bargaining in the state of nature provides a theoretical foundation for the following Hobbesian SWR.

Definition 2 (*Hobbesian SWR*) $p \succcurlyeq^H q$ *if and only if* $\max_{i \in N}(p_i - q_i) \geq \max_{i \in N}(q_i - p_i)$.

It is immediate to note that $\succcurlyeq^H$ is reflexive and complete. In this section, a number of additional properties of $\succcurlyeq^H$ are discussed.

First, $\succcurlyeq^H$ reflects an 'individualistic' logic: there is no aggregation, no distributional consideration: only naked but open bargaining. Agents are all interconnected, but isolated (they can even be imagined as sending public e-mails while isolated in their cubicles). This feature clearly derives from the individualistic logic of the bargaining structure, which seems particularly apt to capture some key characteristics of the interaction between generations.

Second, we stress again that the social preferences incorporated into $\succcurlyeq^H$ do not reflect a notion of justice as impartiality, but rather of justice as mutual advantage. Therefore $p \succcurlyeq^H q$ is properly interpreted as stating that 'p is unobjectionable against q', and $\succcurlyeq^H$ might reflect a minimal notion of 'justice as absence of objections' deriving from the underlying procedure. In other words, although no agent has enough strength to be a dictator, each agent does have veto power. From the viewpoint of intergenerational justice, the structure of bargaining and the resulting SWR establish a relevant form of procedural justice by assigning to each generation the power to *veto* the unpalatable choices of the other generations. This seems a key aspect in current debates on the effect of current decisions on nonrenewable resources and climate change.

Third, although $\succcurlyeq^H$ has very different theoretical foundations as compared to approaches emphasising the notion of justice as impartiality, there are some interesting, and perhaps surprising connections between $\succcurlyeq^H$ and other well-known SWOs. To analyse them, consider the following standard axioms in social choice theory:

Strong Pareto Optimality (SPO): $p > q \Rightarrow p \succ q$.

Anonymity (A): $a \sim \pi a$ for any permutation π.

Cardinality and unit comparability (CU): Let $a_1, a_2, ..., a_n$ be any real numbers and b be any positive number. Then for any $p, q \in \mathcal{R}^n$, $p \succcurlyeq q$ if and only if $(a_1 + bp_1, a_2 + bp_2, ...) \succcurlyeq (a_1 + bq_1, a_2 + bq_2, ...)$.

Next, define the *utilitarian ordering* $\succcurlyeq^U$ by:

$$p \succ^U q \Leftrightarrow \sum_{k=1}^{n} p_k > \sum_{k=1}^{n} q_k.$$

Consider first 2-person societies. If $N = \{1, 2\}$, then the following immediately follows.

7

Proposition 2 *Let $N = \{1, 2\}$. Then $\succcurlyeq^H = \succcurlyeq^U$.*

Note that $\succcurlyeq^H$ is transitive on $\mathcal{R}^2$ and it satisfies SPO, A, and CU. Therefore, the result is implied by D'Aspremont [7], Theorem 3.3.4, p. 51.

Intuitively, (p_1, p_2) is better than (q_1, q_2) if the loss for 1 at p compared to q is smaller than the gain for 2 at p compared to q (or viceversa). Therefore, interestingly, the above argument provides bargaining-theoretic foundations to utilitarianism in $\mathcal{R}^2$ and, if our interpretation of the barganing procedure as a representation of the Hobbesian state of nature is correct, this argument shows a theoretical link between the Hobbesian and the utilitarian traditions.

In more general societies, however, the relation between the Hobbesian SWR and classical utilitarianism is somewhat weaker. If $N = \{1, ..., n\}$, with $n > 2$, rankings can be far from utilitarian: $(2, 2, 2, 2, 2)$ is better than $(3, 3, 3, 3, 0)$. $\succcurlyeq^H$ is not maximin or leximin, either: $(1, 1, 4)$ is better than $(2, 2, 2)$.

Moreover, $\succcurlyeq^H$ satisfies SPO and CU, but in general it does not satisfy A. For example, $(8, 0, 1) \succ^H (0, 1, 8)$. This is not too surprising given that the starting point of our analysis is a notion of justice as mutual advantage and $\succcurlyeq^H$ is the outcome of individualistic bargaining in a Hobbesian state of nature. The axiom of anonymity is a fundamental notion of justice as impartiality, or justice as fairness and it is best embodied in the Rawlsian assumption of individuals acting under a veil of ignorance. Instead in a Hobbesian state of nature agents know their identities.

Interestingly, though, it is immediate to prove that $\succcurlyeq^H$ does satisfy a weaker notion of anonymity (d'Aspremont [7], p. 51):

Weak Anonymity (WA): For all $i, j \in N$, there are $p, q \in \mathcal{R}^n$ such that $p_i > q_i$, $p_j < q_j$, $p_h = q_h$, all $h \neq i, j$, and $q \sim p$.

This highlights some potentially deeper connection with utilitarianism. In fact, let $\lambda \in \mathcal{R}_+^m$ imply $\lambda_i \geq 0$ for all i, and $\lambda \neq \mathbf{0}$, and consider the following generalisation of utilitarianism due to d'Aspremont ([7], p. 46):

Definition 3 (*Generalized m-person utilitarianism*) *A SWO $\succcurlyeq$ is called m-utilitarian, $1 \leq m \leq n$, if for every subset M of m individuals there is some $\lambda \in \mathcal{R}_+^m$ such that for all $p, q \in \mathcal{R}^n$, with $p_h = q_h$ for every h not in M, $p \succcurlyeq q \Leftrightarrow \sum_{i \in M} \lambda_i p_i \geq \sum_{i \in M} \lambda_i q_i$.*

The following result is proved in ([7], Theorem 3.3.4, p.51):

Proposition 3 *A SWO $\succcurlyeq$ is n-utilitarian if and only if it satisfies SPO, WA, and CU.*

Because $\succcurlyeq^H$ satisfies WA, it follows that $\succcurlyeq^H$ is Hobbesian not only in terms of the bargaining-theoretical procedure underlying it, but also in the sense that although the choice of allocations is determined by the welfare of one individual (the Hobbesian sovereign), whose welfare gain is greatest, this individual is not a dictator. Also, the result clarifies formally our previous assertion that there is a deeper connection with utilitarianism. In fact, since $\succcurlyeq^H$ satisfies SPO, WA and CU then it may be said that the only difference with n-utilitarianism is transitivity.

3.2 Discounting

In the case of resource allocation, it is conceivable that ignorance of the technological advances available in the future generates considerable current uncertainty on future prospects. In addition, any evaluation of alternative prospects from the point of view of the 'current' generation may make it difficult to appraise what preferences the generations to come will hold. In this perspective it may be ethically sound to somewhat try and limit the veto power of future generations. To be sure, a number of objections can be, and have been moved to the asymmetric treatment of different generations and the ethical foundations of discounting are at the centre of a vast debate. Our objective here is not to defend discounting and the asymmetric treatment of generations in the allocation of natural resources. Our aim is to highlight the flexibility of the bargaining approach to intergenerational justice and to show that it can accommodate some common intuitions concerning the intergenerational allocation of resources.[3] Theoretically, it may be argued that, unlike in the Rawlsian tradition, an asymmetric treatment of generations is not inconsistent with the construction of the state of nature, given that agents are not placed behind a veil of ignorance and impartiality is not a key requirement of justice. Formally, the bargaining approach can be easily adapted to allow for some limitations on the veto power of future generations. To do so, let w_i denote

---

[3]In the literature on climate change, discounting is used in the authoritative Stern review [16]. For a thorough discussion, see Roemer [14].

the weight associated to generation i, where $w_i \geq 1$ for all i. We can modify the Hobbesian SWR by requiring an allocation p as weakly better than another allocation q whenever for all $e > 0$ and for all i with $q_i - p_i > ew_i$ there exists j with $p_j - q_j > ew_j$. The higher w_i, the lower the weight of generation i in the sense that, for the same e, this generation must have a higher welfare gain in order to object. Using the same manipulations as before, this corresponds to requiring that $\max_{i \in N} \left(\frac{q_i - p_i}{w_i} \right) > e \Rightarrow \max_{i \in N} \left(\frac{p_i - q_i}{w_i} \right) > e$, that is:

Definition 4 (*Discounted Hobbesian SWR*) $p \succcurlyeq^{H_w} q$ *if and only if*

$$\max_{i \in N} \left(\frac{p_i - q_i}{w_i} \right) \geq \max_{i \in N} \left(\frac{q_i - p_i}{w_i} \right).$$

As before, the following is immediate:

Proposition 4 *Let $w \in \mathcal{R}^n$ be a given vector of weights assigned to each generation in N. For any $p, q \in \mathcal{R}^n$, p is unobjectionable against q if and only if $\max_{i \in N} \left(\frac{p_i - q_i}{w_i} \right) \geq \max_{i \in N} \left(\frac{q_i - p_i}{w_i} \right)$*

3.3 Hobbesian social welfare judgments

One great attraction of the welfare criterion $\succcurlyeq^H$ is its applicability to infinite societies: although a Social Welfare Relation arising from intergenerational bargaining cannot be based on any standard bargaining model with an infinite set of players, the advantage of our proposal is that endowing each generation with veto power in an Hobbesian world makes the number of generations irrelevant for the analysis. In this case, the welfare criterion $\succcurlyeq^H$ can be generalised to focus on the supremum of the welfare differences, so that in the intergenerational context with an infinite number of generations $p \succcurlyeq^{H_I} q$ if and only if

$$\sup_{i \in N}(p_i - q_i) \geq \sup_{i \in N}(q_i - p_i).$$

The Hobbesian SWR thus defined on $\mathcal{R}^{\mathcal{N}}$ has various attractive properties in the evaluation of infinite utility streams. For it is reflexive and complete, it satisfies SPO and it does not incur any of the standard problems faced by classical utilitarianism. For example, it provides consistent welfare judgments even if the welfare streams are unbounded. Moreover,

it is possible to prove that the Hobbesian SWR $\succsim^{H_I}$ on $\mathcal{R}^{\mathcal{N}}$ satisfies the property of upper semi-continuity with respect to the sup topology as defined by Hara, Shinotsuka, Suzumura and Xu ([11]).

Putting aside for the moment the specific issues concerning intergenerational justice, does $\succsim^{H}$ lead in general to reasonable social welfare judgments? Consider the sentence: 'If a potential immigrant can increase his utility, by immigrating, more than what any of the citizens of the host country loses, then immigration is justified (independently of the number of citizens in the host country who lose welfare).' The Hobbesian SWR $\succsim^{H}$ precisely captures the intuitions behind this statement and it provides normative foundations for a liberal immigration policy: if no individual in a society can forcefully object to immigration (in the bargaining theoretic sense analysed above), then $\succsim^{H}$ states that the potential immigrant has a right to migrate. This seems intuitively reasonable and normatively appealing, at least if one endorses the individualistic perspective underlying $\succsim^{H}$. Indeed, first-order normative objections against migration very often come from communitarians (see, for example, [17]), who regard supra-individual entities (communities or nations) as the relevant unit of analysis.

A different type of objections can be related to the possible 'extremism' of the SWR, as the following example illustrates.

Example 1 (*Immigration overload*) *Let the utility of an indigenous agent in a country without immigration be α, dropping to $\alpha - \varepsilon$ if there is immigration; normalise to 0 the utility of a prospective immigrant still out of the country, and γ for a new immigrant. Finally assume that an 'old' immigrant allineates perfectly his preferences to the rest of the society (so he would 'suffer' from the presence of further immigrants, and his utility would drop to $\alpha - \varepsilon$). Consider a situation in which there is only one citizen in the home country and there are a large number of potential immigrants outside the country borders. If $(\alpha - \varepsilon, \gamma, 0, 0, 0, ...) \succsim^{H} (\alpha, 0, 0, 0, 0, ...)$, so that $\gamma > \varepsilon$, it also follows that $(\alpha - \varepsilon, \alpha - \varepsilon, \gamma, 0, 0, ...) \succsim^{H} (\alpha, \alpha, 0, 0, 0, ...)$. As a consequence, a society might be quickly 'overrun' by immigration.*

Two counterarguments can be made at this point. First, albeit possibly 'extreme', unlimited migration is not evidently a bad thing. It can be interpreted as a policy of free

movement of people, which liberals (and nonliberals) have long advocated in history. We are not saying that we should advocate free movement, but only that free movement and (potentially) unlimited migration *per se* do not seem to undermine the normative rationale of $\succcurlyeq^H$. Moreover, at a different level of abstraction, one may argue that as more and more people migrate it is unrealistic to keep the benefit (resp. costs) of migration (resp. of receiving immigrants) constant.

Second, even if one rejects the previous argument, extreme cases are arguably not sufficient - per se - to disqualify a SWR. Both the Rawlsian and the utilitarian SWRs, for instance, are vulnerable to 'extreme' counterexamples. Extreme cases are relevant, but they are only part of the story. They can be set aside either if they are unrealistic (but this is a weak reply from a normative viewpoint), or if they are extreme consequences of otherwise desirable principles: for example, the egalitarianism of the maximin is defensible, even though it may be extreme. Therefore the issue is whether the SWR incorporates some relevant normative intuitions - even though it may take them to an extreme. Arguably, at least in the immigration example, the normative intuitions behind $\succcurlyeq^H$ seem reasonable.

Although the prescriptions of $\succcurlyeq^H$ are defensible in the immigration example, there are other scenarios in which they may be less convincing. For example, $\succcurlyeq^H$ could be used to justify a dictator amassing a gigantic wealth provided his increase in welfare is superior to the loss in welfare for *each* of his 1 million subjects. And yet, maybe here the real objection is one of *procedure*, that is dictatorship is undesirable - if the agent is, say, a scientist, why not increase his welfare if nobody loses more than he gains? Also, it is perhaps unrealistic that the dictator amassing gigantic wealth will not impoverish his citizens to the point that their loss of welfare overshadows any pleasure the dictator may enjoy.

It is true, however, that $\succcurlyeq^H$ may have some undesirable implications. Because distributional concerns are irrelevant, it allows for a situation in which some agent (not necessarily a dictator) gets a disproportionately high share of resources provided his gain outweighs the *individual* losses of a huge number of other individuals. The SWR has a strongly individualistic flavour in that it pitches one agent against each and every other agent taken individually. This could imply protection of individuals (against other individuals) but it could also imply primacy of an individual against all others, as in the dictator's example. In

the intergenerational context, this implies that the Hobbesian SWR does not rule out the possibility that one generation gets a disproportionately high share of natural resources. It is worth noting, however, that $\succcurlyeq^H$ does satisfy the axiom of Minimal Equity.[4]

Minimal Equity (ME): For some $p, q \in \mathcal{R}^n$ and $i, j \in N$, $q_i < p_i < p_j < q_j$, $p_h = q_h$, all $h \neq i, j$, and $p \succcurlyeq q$.

Besides, the problem just highlighted is one that the utilitarian welfare ordering also shares. And in a similar vein[5] we can correct the Hobbesian SWR by requiring that any allocation guarantees to each generation a minimal level of utility, so that only if utility exceed such critical level it is deemed valuable. This is a rather natural restriction in the intergenerational context, where the critical utility level can be justified on the basis of sustainability concerns (for a forceful discussion of this notion of sustainability in the context of climate change, see Roemer [14]). So given a vector $c = (c_1, c_2, ...)$ of critical utilities we require any allocation to Pareto dominate the critical allocation. Then we can modify the Hobbesian SWR as:

Definition 5 (*Critical Utility Hobbesian SWR*) $p \succcurlyeq^{H_{cu}} q$ *if and only if* $p \succcurlyeq^H q$ *and* $p > c$.

The problem with $\succcurlyeq^{H_{cu}}$ is that it may solve some problems, but at the potential cost f reducing its effectiveness, because $\succcurlyeq^{H_{cu}}$ is both intransitive and incomplete.

There are additional shortcomings of the Hobbesian SWR which cannot be easily 'fixed': although it has the merit of being a complete relation satisfying the Strong Pareto property, the Hobbesian SWR is not necessarily transitive, as the following examples show.

Example 2 (*Exploiting future generations*) *Suppose there are only three generations, and consider the egalitarian allocation* $(1, 1, 1)$. *An allocation which transferred utility from the second to the first generation would be equally acceptable, since* $(1, 1, 1) \sim^H (2, 0, 1)$, *and this transfer to earlier generations could again be repeated, since* $(2, 0, 1) \sim^H (3, 0, 0)$.

---

[4]See D'Aspremont [7].

[5]See, for example, Blackorby, Bossert and Donaldson [5].

Yet, this latter allocation is strictly preferred to the initial egalitarian one, since $(3, 0, 0) \succ^H$ $(1, 1, 1)$. *So transitivity fails, since*

$$(1, 1, 1) \sim^H (2, 0, 1) \sim^H (3, 0, 0) \ but \ (3, 0, 0) \succ^H (1, 1, 1)$$

This example exploits the fact that since for the Hobbesian SWR the comparison between the utilities if the most and least favoured generation are all that matters for the ranking of any two infinite utility streams, there are many allocations that are indifferent to one another, generating cycles. Unfortunately, disturbing implications of the Hobbesian SWR remain even if we restrict attention to its asymmetric part, as shown in the following example.

Example 3 (*Trading places*) *Suppose there are only three generations, and consider the allocation* $(2, 1, 0)$ *where generations are treated progressively worse the later they are. The third generation could successfully claim* 2 *for itself by decreasing of one unit each the utility of the other two generations, that is* $^H (1, 0, 2) \succ^H (2, 1, 0)$. *Similarly, the second generation could now argue for a similar claim, again decreasing the utility of the other two generations by one unit, since* $^H (0, 2, 1) \succ^H (1, 0, 2)$; *and the first generation could now make exactly the same argument. This however generates the cycle*

$$(2, 1, 0) \succ^H (0, 2, 1) \succ^H (1, 0, 2) \succ^H (2, 1, 0)$$

For a practical application of $\succsim^H$, the existence of 'permutation cycles' of the type above requires some additional mechanism or domain limitation capable of selecting one of the elements of the cycle, or making sure that the presence of such cycles does not preclude the existence of a maximal element. That rules of justice may be cyclical is well-grounded in history: Naeh and Segal [12], for example, argue that some Talmudic rules of justice are deliberately cyclical, and discuss Talmud-inspired ways of breaking the deadlock (or, in some cases, for not doing so).

In the intergenerational context with an infinite number of generations, however, as noted in the introduction, we face a trade off between different principles, and therefore it should not be surprising that given that the Hobbesian SWR is reflexive and complete and it satisfies

the Strong Pareto Principle, and it is upper semi-continuous in the sup norm, transitivity has to be sacrificed.

4   The modified Hobbesian SWR

The cyclical pattern of the previous example would be also ruled out if we tempered this Hobbesian world in which each generation is pitted against every other generation by requiring each generation to exercise its veto power behind an intertemporal veil of ignorance: generation i knows its place in allocation p, but evaluates the alternative vector q in the ignorance of which place it could occupy in this competing allocation. Focusing on societies with a finite number of agents, a modified Hobbesian SWR can be formally defined as follows:

Definition 6 (*Modified Hobbesian SWR*) $p \succcurlyeq^{H^*} q$ *if and only if* $\max\limits_{i,j \in N}(p_i - q_j) \geq \max\limits_{i,j \in N}(q_j - p_i)$.

It is immediate to see that $\succcurlyeq^{H^*} \subset \succcurlyeq^{H}$, and that $\succcurlyeq^{H^*}$ is reflexive and complete. Moreover, we can establish the following:

Proposition 5 *The Modified Social Welfare Relation* $\succcurlyeq^{H^*}$ *is transitive.*

Proof. Let $\bar{p}$ denote the permutation of p such that the components are ranked in ascending order, so that $\bar{p}_1$ is the welfare level of the worst-off agent, $\bar{p}_2$ of the second worst-off agent, and so on. Then $\max\limits_{i,j \in N}(p_i - q_j) \geq \max\limits_{i,j \in N}(q_j - p_i) \Leftrightarrow \bar{p}_n - \bar{q}_1 \geq \bar{q}_n - \bar{p}_1$. Therefore if $p \succcurlyeq^{H^*} q$ and $q \succcurlyeq^{H^*} r$, then $\bar{p}_n - \bar{q}_1 \geq \bar{q}_n - \bar{p}_1$ and $\bar{q}_n - \bar{r}_1 \geq \bar{r}_n - \bar{q}_1$. The latter inequalities imply $\bar{p}_n - \bar{q}_1 + \bar{q}_n - \bar{r}_1 \geq \bar{q}_n - \bar{p}_1 + \bar{r}_n - \bar{q}_1$ or, equivalently, $\bar{p}_n - \bar{r}_1 \geq \bar{r}_n - \bar{p}_1$, as desired.   ∎

Furthermore, it can easily be checked that $\succcurlyeq^{H^*}$ satisfies Anonymity and Minimal Equity, but if $N = \{1, ..., n\}$ with $n > 2$ it does not satisfy Strong Pareto Optimality. Yet it does satisfy the following standard condition.

Weak Pareto Optimality (WPO): $p >> q \Rightarrow p \succ q$.

Furthermore, although in general it does not satisfy CU, it does satisfy the following requirement:

Cardinality and comparability (CC): Let a be any real number and b be any positive number. Then for any $p, q \in \mathcal{R}^n$, $p \succcurlyeq q$ if and only if $a + bp \succcurlyeq a + bq$.

It is again interesting to note that in 2-person societies $\succcurlyeq^{H^*} = \succcurlyeq^U$, and therefore also $\succcurlyeq^{H^*} = \succcurlyeq^H$. Indeed, in 2-person societies it can be shown that $\succcurlyeq^{H^*}$ satisfies Strong Pareto Optimality and CU.

A further property of $\succcurlyeq^{H^*}$ that is worth noting is that it is explicitly representable. Consider the following result:

Proposition 6 *([7], Theorem 3.5.1, p.58) If a SWO $\succcurlyeq$ satisfies WPO and CC, then there exists a numerical function g, homogeneous of degree one, such that for any $p, q \in \mathcal{R}^n$, $\widehat{p} + g(p - \widehat{p}) > \widehat{q} + g(q - \widehat{q})$ implies $p \succ q$, where $\widehat{p} = \frac{1}{n} \sum_{i=1}^{n} p_i$ and $\widehat{q} = \frac{1}{n} \sum_{i=1}^{n} q_i$.*

In the case of $\succcurlyeq^{H^*}$, let $g(x) = \frac{1}{2} \left[\max_{i \in N} x_i + \min_{i \in N} x_i \right]$. Then $\widehat{p} + g(p - \widehat{p}) > \widehat{q} + g(q - \widehat{q})$ if and only if $\widehat{p} + \frac{1}{2} \left[\max_{i \in N} (p_i - \widehat{p}) + \min_{i \in N} (p_i - \widehat{p}) \right] > \widehat{q} + \frac{1}{2} \left[\max_{i \in N} (q_i - \widehat{q}) + \min_{i \in N} (q_i - \widehat{q}) \right]$.

The modified Hobbesian SWO can be seen as mediating between a notion of justice as mutual advantage (see the above discussion of bargaining in the state of nature) and justice as impartiality. This can be seen by noting that, unlike $\succcurlyeq^H$, it does satisfy Anonymity in general.

5   Concluding remarks

In this paper we have studied the idea of analysing the problem of the intergenerational allocation of resources from the standpoint of justice as mutual advantage, rather than the more traditional approach that views justice as impartiality in the treatment of different generations.

This approach has some advantages. First of all, the fact that infinite streams of utilities are the object poses none of the usual difficulties in the analysis. Although a Social Welfare Relation arising from intergenerational bargaining cannot be based on any standard bargaining model with an infinite set of players, the number of generations ceases to play a pivotal role if we assume that each generation can exercise veto power.

Secondly, the Hobbesian approach in which each generation is 'at war' with each other generation seems appropriate in the context of intergenerational bargaining. We are considering a notion of justice as mutual advantage, but as there will never be a situation in which different generations coexist, it seems reasonable to assume that each generation operates in isolation, without forming any coalition with future generations. Indeed, it is easy to conceptualise a future generation objecting to the waste of resources of one of its predecessors. This, however, is also the Achille's heel of the Hobbesian Social Welfare Relation we propose: in its purest version it suffers from intransitivities, while in its modified version it assumes that different generations might take an interest in each other's welfare.

In this respect, the Hobbesian Social Welfare Relation suffers from the usual negative results that we find in the standard literature, pointing to the fact that some shortcomings may be eliminated by restricting the domain, for instance making future utility depend on the current levels at which the resource is depleted. In general, we hope to have enriched the conceptual apparatus by which we deal with the problem of intergenerational allocation of resources. The ultimate validity of the Hobbesian citerion remains an open question, needing much additional research and discussion.

References

[1] Arrow, K. J. (1973) "Rawls's Principle of Just Saving", *Swedish Journal of Economics*, 75: 323-335.

[2] Asheim, G. B. (2010) "Intergenerational equity", *Annual Review of Economics*, 2: 197-222.

[3] Asheim, G. B., W. Bossert, Y. Sprumont, and K. Suzumura (2010) "Infinite-horizon choice functions", *Economic Theory*, 43: 1-21.

[4] Barry, B. (1989) *Theories of Justice*, University of California Press, Berkeley.

[5] Blackorby, C., W. Bossert, and D. Donaldson (1995) "Intertemporal Population Ethics: Critical-Level Utilitarian Principles", *Econometrica*, 63: 1303-1320.

[6] Bossert, W. and K. Suzumura (2008) "Multi-Profile Intergenerational Social Choice", mimeo, Université de Montréal.

[7] d'Aspremont, C. (1985) "Axioms for Social Welfare Orderings"; in Hurwicz, L., D. Schmeidler, and H. Sonnenschein (Eds.) *Social Goals and Social Organization: Essays in Memory of Elisha Pazner*, Cambridge University Press, Cambridge.

[8] Diamond, P. (1965) "The evaluation of infinite utility streams", *Econometrica*, 33: 170–177.

[9] Ferejohn, J. and T. Page (1978) "On the foundations of intertemporal choice", *American Journal of Agricultural Economics*, 60: 269–275.

[10] Gauthier, D. (1986) *Morals by Agreement*, Oxford University Press, Oxford.

[11] Hara, C., T. Shinotsuka, K. Suzumura, and Y. Xu (2008) "Continuity and egalitarianism in the evaluation of infinite utility streams", *Social Choice and Welfare*, 31: 179–191.

[12] Naeh, S. and U. Segal (2009) "The Talmud On Transitivity", mimeo, Boston College.

[13] Roemer, J. E. (1996) *Theories of Distributive Justice*, Harvard University Press, Cambridge.

[14] Roemer, J. E. (2008) "The ethics of distribution in a warming planet", mimeo, Yale University.

[15] Rubinstein, A ., Z. Safra, and W. Thomson (1992) "On the Interpretation of the Nash Bargaining Solution and its Extension to Non-expected Utility Preferences," *Econometrica*, 60: 1171-1186.

[16] Stern, N. (2007) *The Stern Review on the Economics of Climate Change*, Cambridge University Press, Cambridge.

[17] Walzer, M. (1983) *Spheres of Justice*, Basic Books, New York.