TSE working paper series

09-132

Research Group: *Incentives and Public Decision-Making*

December, 2009

The Pareto Principle of Optimal Inequality

ANTOINE BOMMIER AND STÉPHANE ZUBER

Ecole d'économie de Toulouse (TSE) - Manufacture des Tabacs
Aile Jean-Jacques Laffont - 21, allée de Brienne - 31000 TOULOUSE
Tél : +33 (0) 5 61 12 85 89 - Fax : +33 (0) 5 61 12 86 37 - **www.TSE-fr.eu** - contact@TSE-fr.eu

Toulouse School of Economics

The Pareto Principle of Optimal Inequality

Antoine Bommier[1] Stéphane Zuber[2]

Toulouse School of Economics Université Catholique de Louvain

(CNRS, GREMAQ) CORE

Abstract

The Pareto principle is often viewed as a mild requirement compatible with a variety of value judgements. In particular, it is generally thought that it can accommodate different degress of inequality aversion. We show that this is generally not true in time consistent intertemporal models where some uncertainty prevails.

Keywords: Inequality aversion, Pareto Principle, Uncertainty, Time Consistency.

JEL Classification numbers: D6, D7, D81, D9.

[1]Université des Sciences Sociales - Manufacture des Tabacs, 21 allée de Brienne, 31000 Toulouse, France. E-mail: Antoine.Bommier@univ-tlse1.fr.

[2]CORE, Université Catholique de Louvain, Voie du Roman Pays 34, 1348 Louvain-La-Neuve, Belgium. E-mail: Stephane.Zuber@uclouvain.be.

1 Introduction

It has long been recognized that adherence to the Pareto principle may have distributional consequences. In particular, the set of Pareto optimal allocations may not contain any egalitarian allocation, so that equality and optimality are impossible to reconcile. Philosophers, especially those adhering to the egalitarian tradition, have intensively debated as to whether inequalities should be permitted in such cases, or the Pareto principle abandoned.

Although there is no possible dispute as to the fact that the Pareto principle may constrain the Social Planner in his willingness to achieve equality, it is also widely believed that adherence to the Pareto principle does not make all concerns for inequality irrelevant. The set of Pareto optimal allocations is typically thought of as being rich enough for the planner to be left with a non trivial choice where his preferences for equality may matter. Stated differently, it is usually believed that accepting the Pareto principle does not prevent diverse degrees of inequality aversion from being expressed.

This conventional wisdom about the limited restrictiveness of the Pareto principle is perfectly correct in deterministic or static cases. The aim of this paper is to show that it does not extend to realistic cases, where dynamics and uncertainty are inevitably at play. We show that in such cases adhering to the Pareto principle generally leaves practically no room for redistributive considerations. More precisely, we prove that, except in particular cases, two Paretian social observers with consequentialist and time consistent preferences cannot express different but comparable concerns for equality. Given a social observer with Paretian, consequentialist and time consistent preferences, it is impossible to find another Paretian, consequentialist and time consistent social observer who would exhibit more (or less) aversion to inequality.

To grasp the intuition behind the result, imagine a population of identical individuals. In period one, all these individuals face the same decision problem, which consists of making a choice under uncertainty, knowing that the consequence of one's choice will have no impact on the others' well being. All individuals being identical they are willing to take identical decisions. Since each individual decision has no impact on the others, a Paretian social planner has no other possibility than deciding for everybody what they would have decided for themselves. In particular the planner's decision is independent of whether individuals' fates are determined by a single lottery that apply to all, or by independent individual lotteries. These two cases would however generate very different degrees of inequalities in period 2. But the planner, being Paretian, has to accept the levels of inequality arising from individuals' preferred choices. He is thus left with no possibility to express his own judgement about the socially optimal level of inequality: the acceptance of the Pareto principle ex-ante prevents inequality considerations ex-post.

Alternatively, one grasp the intuition of the result by considering Diamond (1967)'s criticism of Harsanyi's (1955) theorem. A simple way to obtain time consistency in a dynamic setting involves relying on the expected utility framework. Harsanyi (1955) proved that the only way to get a Paretian aggregation of preferences within the expected utility framework consists in having a social utility function which is an affine combination of individuals' utility functions. Diamond (1967) argued that this form of preference aggregation annihilates the possibility of redistributing utility.

Diamond's remark, as well as the simple example we suggested above, motivate our research. Our aim is to show that, far form being confined to particular cases (as in the first example), or confined to a particular setting (expected utility framework, in the Diamond-Harsanyi controversy), these insightful examples are the mere expression of a general impossibility result whose formalization is precisely the object of our paper. We will show indeed that considerations of inequality aversion become irrelevant when the planner adheres to the Pareto principle while his dynamic preferences satisfy some basic conditions - namely time consistency and consequentialism. While egalitarians pointed at the Pareto principle as a potential obstacle to reach full equality, the Pareto principle has much stronger consequences: it basically forbids the expression of different concern for equality.

The present paper relates two lines of research: inequality aversion, on the one hand, and dynamic choice under uncertainty, on the other hand. As such it uses contributions from two broad but almost disconnected fields of literature. We borrow from the works by Yaari (1969), Kihlstrom and Mirman (1974), Jewitt (1989) Grant and Quiggin (2005) and Bosmans (2007) to provide a general definition of comparative inequality aversion.[1] Concerning dynamic choice under uncertainty, we use the huge literature on the link between consequentialism, linearity in mixture and time consistency. Among the most relevant references are Myerson (1981), Johnsen and Donaldson (1985) and Sarin and Wakker (1994). Two important conclusions of this literature are the following: first, linearity in mixture is a very usual and practical way of ensuring the time consistency of choices (Myerson, 1981). It is for this reason that one part of our paper will focus

[1]Note that all these papers except that of Bosmans deal with risk or uncertainty aversion. There is however a direct link with the concept of inequality aversion we will use in the paper.

on preferences that are linear in mixture. Second, in settings which are not embedded with a natural mixture operation (e.g. when considering a finite set for the state of the world), the assumption of time consistency is -in a consequentialist approach- akin to a separability assumption which gives a particular structure to the representation of dynamic preferences (Johnsen and Donaldson, 1985). Consequentialism and time consistency are in fact at the heart of the folding back technique which is almost universally adopted in recursive choices (Sarin and Wakker, 1994).

The remainder of the paper is structured as follows. In Section 2, we introduce the setting and provide a general definition of inequality aversion. In Section 3, we consider the possibility of different degrees of inequality aversion when social observers are Paretian. We first consider the case where they have preferences that are linear in mixture. We show that it is impossible to have two Paretian social observers with one being more inequality averse than the other. We then study a more general set up where the choice space does not permit mixtures. Assuming time consistency of choices, we obtain a similar result: a Paretian social observer cannot be more inequality averse than another one after the initial period. Section 4 concludes.

2 Inequality aversion

Since the pioneering works by Atkinson (1970) and Kolm (1976), the degree of inequality aversion has been identified as a key element for policy assessment. Several papers have provided characterizations of the degree of inequality aversion for specific models. Atkinson (1970) and Kolm (1976) are the basic references in the case of additive separable social evaluation functions. Donaldson and Weymark (1980) derive characterizations in the case of Yaari social evaluation functions (which generalize the Gini evaluation function).

However, in order to obtain general results, comparison of inequality aversion should be independent of the acceptance of a particular model of social choice. Moreover, a general definition should not be contingent on a particular structure for the set of social consequences. In particular, to be relevant in complex settings (as with the case of dynamic choice under uncertainty) it is necessary to consider a more general problem than the allocation of a single transferable commodity between individuals.

Similar considerations are found in the literature on choice under uncertainty. While the seminal contributions of Pratt (1964) and Arrow (1965) were restricted to von Neumann and Morgenstern (vNM) preferences over the set of monetary lotteries, it soon appeared necessary to have a definition of comparative risk aversion that would neither rely on the expected utility framework, nor be restricted to the case of unidimensional lotteries. The contributions by Yaari (1969), Kihlstrom and Mirman (1974), Jewitt (1989) and Grant and Quiggin (2005) - in the case of uncertainty aversion - worked in that direction.

The common principle in these contributions is to associate a relation "more risk averse than" (which involves comparing preferences) with any relation "riskier than" (which involves comparing random elements). More precisely, given a relation "riskier than" a decision maker A is said to be less (or no more) risk averse than B if any x which is riskier than y and preferred to y by B is also preferred to y by A. The advantage is that such a definition does not involve assuming a particular model of choice under uncertainty, and that it can be applied to complex domains of consequences.

Transposition to social choice for comparisons of inequality aversion is explicit in Bosmans (2007). The same route will be taken in the present paper. However, a difference will be that - for reasons explained below - we shall consider the relation "more unequal than" introduced by Hammond (1976). This relation, which was not considered by Bosmans, corresponds to the notion of single crossing of the cumulative distribution functions introduced by Jewitt (1989).

2.1 The setting

We consider society composed of two individuals, denoted 1 and 2. The set of all conceivable (or feasible) social alternatives is denoted by X. It is a general set with no particular structure. We assume that each individual $i \in \{1, 2\}$ is endowed with a preference relation $\succeq_i$ on X.

We aim at comparing social preferences over X in terms of inequality aversion. Discussing matters related to inequality involves being able to (ordinally) compare the well being of the two individuals. Indeed, it would be meaningless to pretend taking into account the inequality of a social alternative, if one is not even able to tell which of the two individuals is better off.

We will therefore define a Social Observer ("SO", hereafter) as being a utility function defined on X, allowing social alternatives to be ranked, and an interpersonal comparison function allowing the well-beings of individuals in the society to be compared. More precisely:

Definition 1 *A Social Observer is the combination of:*

1. *A social evaluation function* $W : X \to \mathbb{R}$.

2. *An interpersonal utility function* $U : X \times \{1, 2\} \to \mathbb{R}$.

The social evaluation function W orders social alternatives. It represents the preferences of the SO. The interpersonal utility function compares the relative situations of the two individuals in alternative states of affairs. The inequality $U(x, i) \geq U(\hat{x}, j)$ denotes the ethical judgement that individual i is better-off in the social alternative x than j in the social alternative $\hat{x}$.

In most of the literature on inequality measurement, the interpersonal utility function is not formalized. For instance, papers on income inequality typically do not refer to such an interpersonal utility functions. Still they assume that transfers from the rich to the poor is socially desirable, which makes sense only if it is implicitly assumed that an individual who earns more than another one is also better off.

In our general setting, we need a way to compare individual's welfare, so that it becomes possible to tell whether a policy that increase the welfare of an individual but decrease the welfare of the other enhances or reduces inequality. This welfare comparison is provided by the interpersonal utility function. Doing so, we also follow a direction very common in the literature on multidimensional inequality. It consists in first aggregating attributes at the individual level and then define inequality in terms of utility.[2]

[2]See Atkinson and Bourguignon (1982) and Maasoumi (1986).

Since we have defined SOs as being utility functions (and not preference relations) it will be useful to have the following definition:

Definition 2 *Two functions* $V^A : K \rightarrow \mathbb{R}$ *and* $V^B : K \rightarrow \mathbb{R}$ *are* ordinally congruent *if there exists an increasing function* $\phi : \mathbb{R} \rightarrow \mathbb{R}$ *such that*

$$V^A = \phi(V^B)$$

The above definition can apply either to SOs' social evaluation functions (when $K = X$) or to their interpersonal utility functions (when $K = X \times \{1, 2\}$). Two SOs with ordinally congruent social evaluation functions have the same preferences over social alternatives. Two SOs with ordinally congruent interpersonal evaluation functions always concur when comparing the individuals' welfare.

If we endorse individualistic views, restrictions have to be placed upon the social evaluation function and the interpersonal utility function. The SO's assessment of individual i's welfare must conform with i's point of view. Moreover, the social evaluation function must accord with the Pareto principle which states that the unanimous agreement of individuals has to be respected. We will call an SO that satisfies these two properties 'Paretian'.

Definition 3 *An SO* (W, U) *is* Paretian *if:*

1. *His interpersonal utility function is such that* $U(., i)$ *represents the preferences of individual i:*

$$U(x, i) \geq U(\hat{x}, i) \Longleftrightarrow x \succeq_i \hat{x} \qquad \forall x, \hat{x} \in X, \;\; \forall i = 1, 2$$

2. *There exists a strictly increasing function $\mathcal{W}$ such that the social evaluation function $W(x)$ is given by:*

$$W(x) = \mathcal{W}\Big(U(x,1), U(x,2)\Big) \tag{1}$$

2.2 Comparative inequality aversion

In order to implement the procedure by Bosmans (2007) for defining a relation "more inequality averse than" one first needs to define a relation "more unequal than". Bosmans explored three definitions. One is the M-concept where x is at least as unequal as $\hat{x}$ if and only if $\hat{x}$ describes a situation with perfect equality. The two other concepts are based respectively on Lorenz domination and on relative differentials quasi-ordering, both applied to the distribution of the $U(., i)$.

None of the three definitions seemed adequate to us. The M-concept is, as its name suggests, extremely minimalist, and leads to a very weak notion of comparative inequality aversion. In practice, the M-concept affords interesting results only when combined with additional assumptions on the form of the social evaluation function (e.g. by assuming additive separability). The two other concepts are more constraining, but have the drawbacks of not being ordinal, in the sense that two SOs with ordinally congruent interpersonal utility functions would not necessarily agree on inequality comparisons.

Rather, we prefer to use the notion introduced by Hammond (1976) which allows social alternatives to be compared even if none of them is characterized by full equality.

Definition 4 *Given an interpersonal evaluation function U, a social alternative $x \in$*

X is said to be at least as unequal as $\hat{x} \in X$*, denoted* $xI_U\hat{x}$*, if, for a pair* $(i, j) \in \{(1,2),(2,1)\}$*:*

$$U(x,i) \leq U(\hat{x},i) \leq U(\hat{x},j) \leq U(x,j)$$

or

$$U(x,i) \leq U(\hat{x},j) \leq U(\hat{x},i) \leq U(x,j)$$

In addition, $x \in X$ *is* more unequal than $\hat{x} \in X$*, denoted* $x\widehat{I}_U\hat{x}$*, if there is at least one strict inequality between the first and second terms or the third and fourth terms in the above inequations.*

It is clear that the relations I_U and $\widehat{I}_U$ are transitive. Moreover, I_U and $\widehat{I}_U$ are ordinal concepts in the sense that if $V = \phi(U)$ for an increasing function ϕ then $I_U = I_V$ and $\widehat{I}_U = \widehat{I}_V$.

We can now implement Bosmans's procedure to obtain a relation of comparative inequality aversion.

Definition 5 *An SO* (W^A, U^A) *is* at least as inequality averse as *an SO* (W^B, U^B)*, if and only if, for any* $\hat{x} \in X$*:*

$$\left\{x \in X : x\widehat{I}_{U^A}\,\hat{x} \;\; \textit{and} \;\; W^A(x) \geq W^A(\hat{x})\right\} \subset \left\{x \in X : x\widehat{I}_{U^B}\,\hat{x} \;\; \textit{and} \;\; W^B(x) \geq W^B(\hat{x})\right\}$$

From the definition it is possible to derive characterization results in specific settings (additively separable, Yaari or Quiggin social welfare functions, and so forth).[3] But the great advantage of Definition 3 is that it does not depend on a specific model of social

[3]For more details in a unidimensional setting, see Grant and Quiggin (2005).

evaluation, in the same way as Yaari's (1969) definition is not restricted to the expected utility model. Definition 3 is consistent with but more general than views expressed in the literature on inequality measurement.

Our very general definition of inequality aversion can be applied to complicated consequence domains, for instance dynamic settings involving uncertainty. We are therefore able to compare SO's inequality aversion in these settings.

3 Time Consistency, consequentialism and Inequality Aversion

There is a broad literature on what should be the desirable properties of an SO's preferences in a dynamic context. A common assumption is that of time consistency, which states that if an SO thinks that he will take a given action under certain circumstances, he will actually stick to his plan if these circumstances do occur. Another common assumption, central in the utilitarian doctrine, is that the SO should be consequentialist, in the sense that his preferences regarding the future should be independent of what could have happened in the past in other circumstances.

The fact that a well-behaved SO should be time consistent is not very controversial. The assumption of consequentialism has however been criticized in several instances, most notably by Diamond (1967), Machina (1989) and Epstein and Segal (1992). Yet, in dynamic settings, consequentialism remains endorsed by most theorists. Defenses of consequentialism have been proposed by Hammond (1988), Sarin and Wakker (1994), and to some extend by Fleurbaey (2007).

It is not our purpose, in the present paper, to debate about the desirability of assuming time consistency and consequentialism. Our purpose is rather to emphasize that these assumptions combined with the Pareto principle turn out to have significant implications in terms of inequality aversion.

We will consider two different settings. In Section 3.1 we will explore the case of preferences that are linear in mixture. Several papers, including Hammond (1988), or Myerson (1981), argue that this property is necessary to guarantee time consistency. In the context of consequentialist sequential choices, several papers have actually shown that time consistency and the independence condition which corresponds to the linearity in mixture assumption are closely related. More precisely, it has been shown in consequentialist frameworks that time consistency combined with some other property implies either the independence condition or the expected utility model.[4]

Moreover, the wide majority of choice models under risk and/or uncertainty fulfill an assumption of (possibly partial) linearity in mixture. Examples are the expected utility, the subjective expected utility *à la* Anscombe and Auman (1963) or, in a restricted sense, the dynamic model by Kreps and Porteus (1978). More recent models[5] also satisfy a form of partial linearity of mixture, and therefore will fit in this part. We will indeed show that partial linearity in mixture makes it possible to derive a general impossibility result about inequality aversion.

Still, especially when considering a relatively poor information context (like a finite

[4]The other property can be any one of the following: 1/ timing indifference (Chew and Epstein, 1989); 2/ reduction of compounded lotteries (Karni and Schmeidler, 1991); 3/ interchangeability of consecutive decision nodes (Sarin and Wakker, 1994).

[5]For instance the Maxmin Expected Utility model of Gilboa and Schmeidler (1989), the Choquet Expected Utility of Schmeidler (1989) and more generally the invariant biseparable preferences of Ghirardato, Maccheroni and Marinacci (2004)

set for the state of the world), it is known since Johnsen and Donaldson (1985) that linearity in mixture is not required to obtain time consistency. In Section 3.2, we will therefore follow the approach of Johnsen and Donaldson and show that a similar impossibility result can be derived.

3.1 Linearity in mixture

In this section, the temporal dimension of the decision problem need not be made explicit, since only the assumption of linearity in mixture will be needed for our results. We will simply assume agents have preferences over an abstract set X on which an operation of mixture is defined. To apply our result to particular settings, such as situations of risk, uncertainty, or temporal risk *à la* Kreps and Porteus (1978), one only needs to consider the appropriate set X.

A mixture operation $\oplus$ on X is an operation, which for all x and $\hat{x}$ belonging to X and all $\alpha \in [0,1]$, provides an element $\alpha x \oplus (1-\alpha)\hat{x} \in X$ such that:

$$\begin{cases} 0x \oplus 1\hat{x} = \hat{x}, \\ \alpha x \oplus (1-\alpha)\hat{x} = (1-\alpha)\hat{x} \oplus \alpha x \quad \text{for all } x, \hat{x} \in X, \\ \alpha[\beta x \oplus (1-\beta)\hat{x}] \oplus (1-\alpha)\hat{x} = \alpha\beta x \oplus (1-\alpha\beta)\hat{x} \quad \text{for all } \alpha, \beta \in [0,1] \end{cases}$$

One typical example of mixture space is the space of all simple lotteries over a set of payoffs. Let L_1 and L_2 be two lotteries. Let L_3 be the compound lottery consisting of playing L_1 with probability α and L_2 with probability $1-\alpha$. The lottery L_3 is a mixture of L_1 and L_2, $L_3 = \alpha L_1 \oplus (1-\alpha)L_2$.

Definition 6 *A function* $F : X \to \mathbb{R}$ *is* continuous with respect to the mixture operation $\oplus$ *if, for all* $x, \hat{x} \in X$*, the function* $\psi_{x,\hat{x}} : [0,1] \to \mathbb{R}$ *such that* $\psi_{x,\hat{x}}(\alpha) = F\big(\alpha x \oplus (1-\alpha)\tilde{x}\big)$*, is continuous.*

Definition 7 *A SO* (W, U) *is* continuously linear in the mixture operation $\oplus$ *with respect to the subset* $\widetilde{X} \subset X$ *if and only if, for all* $x, \hat{x} \in X$*,* $\tilde{x} \in \widetilde{X}$*,* $\alpha \in (0,1)$ *and* $i \in \{1, 2\}$

$$\begin{aligned} W(x) &\geq W(\hat{x}) \Longleftrightarrow W\Big(\alpha x \oplus (1-\alpha)\tilde{x}\Big) \geq W\Big(\alpha \hat{x} \oplus (1-\alpha)\tilde{x}\Big) \\ U(x,i) &\geq U(\hat{x},i) \Longleftrightarrow U\Big(\alpha x \oplus (1-\alpha)\tilde{x}, i\Big) \geq U\Big(\alpha \hat{x} \oplus (1-\alpha)\tilde{x}, i\Big) \end{aligned}$$

and the three functions W*,* $U(.,1)$ *and* $U(.,2)$ *are continuous with respect to the mixture operation* $\oplus$*.*

Whenever $\widetilde{X} = X$ in Definition 7, we have the usual linearity in mixture property that is satisfied by the expected utility representation, or Anscombe and Aumann (1963) subjective expected utility. But Definition 7 can also encompass many other cases. For instance, when discussing preferences over temporal lotteries, the Kreps and Porteus (1978) recursive expected utility model satisfy linearity in mixture with respect to the subset of temporal lotteries resolving in the last period of time.

Definition 7 also encompasses the property of certainty independence if $\widetilde{X}$ is the set of constant acts in an Anscombe-Aumann framework. Certainty independence is used to characterize decision models that generalize subjective expected utility.[6] Our result thus extends to these models of choice and their dynamic extensions.

[6] See Ghirardato, Maccheroni and Marinacci (2004).

We also need to introduce the following definition:

Definition 8 *Given a SO* (W, U)*, a subset* $\widetilde{X} \subset X$ *is said to be* socially ambivalent *if there exist* $\tilde{x}_1$ *and* $\tilde{x}_2 \in \widetilde{X}$ *such that*

$$U(\tilde{x}_1, 1) < U(\tilde{x}_1, 2) \quad \text{and} \quad U(\tilde{x}_2, 1) > U(\tilde{x}_2, 2)$$

Social ambivalence is a condition on the utility possibilities set. It amounts to assuming that individuals' welfare ranks can be inverted, so that one individual is not systematically better off from the SO's point of view. Social ambivalence is an ordinal condition since two SOs with ordinally congruent interpersonal utility functions will always agree on whether a subset is socially ambivalent.

Now we can state our result:

Proposition 1 *Consider two Paretian SOs* (W^A, U^A) *and* (W^B, U^B)*, who are continuously linear in mixture with respect to a subset* $\tilde{X}$ *which is socially ambivalent for A. If A is at least as inequality averse as B, then A and B have ordinally congruent social evaluation functions.*

Proof. Assume that $W^A(x) \geq W^A(\hat{x})$. We need to show that $W^B(x) \geq W^B(\hat{x})$.

If x Pareto dominates $\hat{x}$ (in the broad sense that includes Pareto indifference) then $W^B(x) \geq W^B(\hat{x})$. So consider the case where x does not Pareto dominate $\hat{x}$ and is not Pareto indifferent to it. Since $W^A(x) \geq W^A(\hat{x})$ we can assume without loss of generality that[7]

$$U^A(x,1) < U^A(\hat{x},1)$$
$$U^A(\hat{x},2) < U^A(x,2)$$

Now consider the mixtures

$$x' = \alpha x \oplus (1-\alpha)\tilde{x}_2$$

and

$$\hat{x}' = \alpha \hat{x} \oplus (1-\alpha)\tilde{x}_2$$

where $\tilde{x}_2$ is the social alternative described in Definition 8.

Since $U^A(\tilde{x}_2, 2) > U^A(\tilde{x}_2, 1)$, and because $U^A(., 1)$ and $U^A(., 2)$ are continuous with respect to $\oplus$, we have for α close enough to 0

$$U^A(\hat{x}', 1) \leq U^A(\hat{x}', 2)$$

By linearity in mixture, we also have $U^A(\hat{x}', 2) < U^A(x', 2)$ and $U^A(x', 1) < U^A(\hat{x}', 1)$. Thus, for α close enough to 0,

$$U^A(x', 1) < U^A(\hat{x}', 1) \leq U^A(\hat{x}', 2) \leq U^A(x', 2)$$

[7] The case $U^A(\hat{x}, 1) < U^A(x, 1)$ and $U^A(x, 2) < U^A(\hat{x}, 2)$ can be dealt with using the same methods as below. The only difference is that we should then use $\tilde{x}_1$ rather than $\tilde{x}_2$ for defining x' and $\hat{x}'$ in the remainder of the proof.

Hence $x' \hat{I}_{U^A} \hat{x}'$. By linearity in mixture, it is also the case that $W^A(x') \geq W^A(\hat{x}')$. Thus, by comparative inequality aversion, $W^B(x') \geq W^B(\hat{x}')$. Finally, the linearity in mixture of B's preferences implies that $W^B(x) \geq W^B(\hat{x})$. ■

Proposition 1 formalizes and generalizes Diamond's (1967) criticism of Harsanyi's theorem. Diamond asserted that Harsanyi's criterion is unfair because it precludes redistributions of welfare. The argument was based on particular cardinal utility functions for the individuals. Proposition 1 does not make such assumptions: only ordinal representations are used. Besides, Proposition 1 clarifies the argument by showing that inequality aversion is fixed when comparable. Proposition 1 is also not confined to a particular choice situation; it encompasses most frameworks of choice under uncertainty.

One crucial assumption to obtain the result is that the subset $\widetilde{X}$ is socially ambivalent. The following example illustrates that, when the assumption does not hold, it is possible for a Paretian SO to be more inequality averse than another.

> **Example:** Consider two SOs (W^A, U) and (W^B, U) using the same interpersonal utility function U. A has a social evaluation function $W^A(x) = \alpha_1^A U(x,1) + \alpha_2^A U(x,2)$ and B has a social evaluation function $W^B(x) = \alpha_1^B U(x,1) + \alpha_2^B U(x,2)$, where $\alpha_1^A + \alpha_2^A = \alpha_1^B + \alpha_2^B = 1$. Assume also that $U(x,1) < U(x,2)$ for any $x \in X$, so that X is not socially ambivalent.
>
> In that case, if $\alpha_1^A > \alpha_1^B$, then A is more inequality averse than B. Indeed, for any $x \in X$, any more unequal situation $\hat{x}$ is such that $U(\hat{x},1) \leq U(x,1) \leq U(x,2) \leq U(\hat{x},2)$. If $W^A(\hat{x}) \geq W^A(x)$, then $\alpha_1^A U(\hat{x},1) + \alpha_2^A U(\hat{x},2) \geq \alpha_1^A U(x,1) + \alpha_2^A U(x,2)$. Consequently, $\frac{U(\hat{x},2)-U(x,2)}{U(x,1)-U(\hat{x},1)} \geq \frac{\alpha_1^A}{\alpha_2^A} > \frac{\alpha_1^B}{\alpha_2^B}$,

because $\alpha_1^A > \alpha_1^B$ and $\alpha_1^A + \alpha_2^A = \alpha_1^B + \alpha_2^B = 1$. We hence find that $\left(\hat{x}\widehat{I}_U x \text{ and } W^A(\hat{x}) \geq W^A(x)\right) \Longrightarrow W^B(\hat{x}) > W^B(x)$.

The conclusion of this section is that, whenever a subset of X is socially ambivalent, we must weaken either the Pareto principle or (restricted) linearity in mixture in order to escape the negative result of Proposition 1.

In Harsanyi's framework, it has been suggested that we should weaken linearity in mixture (Diamond, 1967). Some attempts in this direction have actually been made (Epstein and Segal, 1992; Grant, Kajii, Polak and Safra, 2006). A difficulty with these approaches is that they are prone to induce time inconsistent judgments under consequentialist principles. The next section indeed shows that time consistency combined with consequentialism and the Pareto principle prohibits comparisons in terms of inequality aversion.

3.2 A dynamic framework

Most of the time consistent dynamic models that are found in the literature fulfill an assumption of linearity of mixture. The result of Proposition 1 can then be directly applied. Though, linearity in mixture is not necessary to ensure time consistency. Indeed, as explained in Johnsen and Donaldson (1985), under consequentialist principles, time consistency implies a particular separability condition that they called "conditional weak independence". This condition is much weaker than the separability assumption underlying the expected utility model.

The aim of this section, is to show that a result similar to Proposition 1 can be obtained, when considering the time consistency condition formalized by Johnsen and Donaldson. Hence, while no linearity assumption is made, the conclusion of Proposition 1 still holds. This therefore shows the generality of our findings.

We consider a simple consequentialist intertemporal setting that extends to T periods the model proposed by Johnsen and Donaldson (1985). The time horizon is finite and each period denoted by $t \in \{1, \cdots, T\}$. Each period, any of S states of the world can arise. Let $\Sigma = \{1, \cdots, S\}$. For any arbitrary set K, let $\mathcal{F}(K)$ be the set of acts defined as follows: $\mathcal{F}(K) = \{f : \Sigma \to K\}$. In period t, the possible outcomes belong to the set Z_t. Each period, the decision maker must choose a course of action. We describe the intertemporal choice problem recursively. We define the sets of temporal acts as $Y_T = Z_T$ and, for all $t \in \{1, \cdots, T-1\}$, $Y_t = Z_t \times \mathcal{F}(Y_{t+1})$. For $t > 1$, we define histories h_t as the collection of past outcomes: $h_t \in \mathcal{H}_t = \prod_{\tau=1}^{t-1} Z_\tau$. In a period t and given a history h_t, the decision maker is able to choose a course of action using the function $U_{h_t,t} : Y_t \to \mathbb{R}$. The decision maker's dynamic preferences are completely represented by the collection $\Big\{U_{h_t,t} : t = 1, \cdots, T; h_t \in \mathcal{H}_t, \Big\}$. We call $\Big\{U_{h_t,t} : t = 1, \cdots, T; h_t \in \mathcal{H}_t, \Big\}$ - in short $\{U_{h_t,t}\}$ - a process of preferences.

This formalization implicitly assumes that the decision maker is *consequentialist*, in the following usual sense: what might have happened in unrealized states of the world do not impact preferences over the future. We also assume that preferences are state independent: preferences in period t do not dependent on which states of world have occurred until period t included.

In line with Johnsen and Donaldson (1985), we consider the following additional property of processes of preferences:

Definition 9 *A process of preferences* $\{U_{h_t,t}\}$ *is* time consistent *if, for any* $t \in \{1, \cdots, T-1\}$*,* $h_t \in \mathcal{H}_t,\ z_t \in Z_t,\ f_t, \hat{f}_t \in \mathcal{F}(Y_{t+1})$*:*

$$U_{(h_t,z_t),t+1}\Big(f_t(s)\Big) \geq U_{(h_t,z_t),t+1}\Big(\hat{f}_t(s)\Big) \quad \forall s \in \Sigma \Longrightarrow U_{h_t,t}\Big(z_t, f_t\Big) \geq U_{h_t,t}\Big(z_t, \hat{f}_t\Big)$$

If, furthermore, $U_{(h_t,z_t),t+1}\Big(f_t(s')\Big) > U_{(h_t,z_t),t+1}\Big(\hat{f}_t(s')\Big)$ *for some* $s' \in \Sigma$ *then:* $U_{h_t,t}\Big(z_t, f_t\Big) > U_{h_t,t}\Big(z_t, \hat{f}_t\Big)$*.*

In the current framework, there is no definition of the mixture operation that could render linearity in mixture appealing. The first period outcome is certain and cannot be randomized. But, as emphasized by Johnsen and Donaldson (1985), the above time-consistency principle implies "conditional weak independence", a notion formally defined in their paper. Using results by Koopmans (1972), Johnsen and Donaldson establish the utility representations that are equivalent to that separability condition, emphasizing that these are consistent with non linear models. In particular, the expected utility model is a special case that corresponds to a stronger separability condition (strong independence).

Now, consider two individuals 1 and 2, each endowed with a process of preferences $\{U^i_{h_t,t}\}$. In the dynamic framework we have set up, we need to enlarge the notion of an SO to include dynamic preferences.

Definition 10 *An Social Observer is the combination of:*

1. *A collection of social evaluation functions* $W_{h_t,t} : Y_t \to \mathbb{R}$*, one for each* $t \in \{1, \cdots, T\}$ *and* $h_t \in \mathcal{H}_t$*.*

2. *A collection of interpersonal utility functions* $U_{h_t,t} : X \times \{1,2\} \to \mathbb{R}$ *one for each* $t \in \{1, \cdots, T\}$ *and* $h_t \in \mathcal{H}_t$.

An SO is *Paretian* if the definition of a Paretian SO provided in Definition 3 applies to $W_{h_t,t}$ and $U_{h_t,t}$ for each $t \in \{1, \cdots, T\}$ and $h_t \in \mathcal{H}_t$. A SO $\{(W_{h_t,t}, U_{h_t,t})\}$ is *time consistent* if each process of preferences $W_{h_t,t}$, $U_{h_t,t}(.,1)$ and $U_{h_t,t}(.,2)$ is time consistent.

In the present framework, $X \equiv Y_1$. Comparative inequality aversion is expressed in terms of first period social evaluation functions and interpersonal utility functions.

As above, an additional assumption must be made to obtain the analog of Proposition 1. The assumption guarantees that individuals' welfare rank can be inverted by ensuring a particular outcome in sufficiently numerous states of the world.

Definition 11 *A state of the world* $s \in \Sigma$ *is* socially revertible *for an SO* $\{(W_{h_t,t}, U_{h_t,t})\}$ *if, for all* $z_1 \in Z_1$*, there exist* $\widetilde{y}_2^1$ *and* $\widetilde{y}_2^2$ *in* Y_2 *such that, whatever* $y_2 \in Y_2$*, if* $f_1(s) = \hat{f}_1(s) = y_2$*, and, for all* $s' \neq s$*,* $f_1(s') = \widetilde{y}_2^1$ *and* $\hat{f}_1(s') = \widetilde{y}_2^2$ *then:* $U_1\Big((z_1, f_1), 1\Big) > U_1\Big((z_1, f_1), 2\Big)$ *and* $U_1\Big((z_1, \hat{f}_1), 2\Big) > U_1\Big((z_1, \hat{f}_1), 1\Big)$.

The existence of a socially revertible state of the world represents two ideas. First, this state of the world is sufficiently unimportant ('sufficiently unlikely') so that what occurs in other states of the world can fully fix the relative situation of individuals. Second, it is possible to invert individuals' welfare ranks. The second idea is closely related to the existence of a socially ambivalent set of social alternatives.

Whenever there exists a socially invertible state of the world, we can obtain a result similar to Proposition 1.

Proposition 2 *Consider two Paretian SOs* $\{(W^A_{h_t,t}, U^A_{h_t,t})\}$ *and* $\{(W^B_{h_t,t}, U^B_{h_t,t})\}$*, who are time consistent. If 1/ there exists a state of the world* $s \in \Sigma$ *that is socially invertible for* $\{(W^A_{h_t,t}, U^A_{h_t,t})\}$*, and 2/ A is at least as inequality averse as B; then A and B have ordinally congruent social evaluation functions after the initial period.*

Proof. We first prove that A and B must have the same preferences in period 2. Consider any $z_1 \in Z_1$, y_2 and $\hat{y}_2$ in Y_2. Assume that $W^A_{z_1,2}(y_2) \geq W^A_{z_1,2}(\hat{y}_2)$. We want to show that $W^B_{z_1,2}(y_2) \geq W^B_{z_1,2}(\hat{y}_2)$.

If y_2 Pareto dominates $\hat{y}_2$ given z_1, then $W^B_{z_1,2}(y_2) \geq W^B_{z_1,2}(\hat{y}_2)$. So consider the case where y_2 does not Pareto dominate $\hat{y}_2$. Assume for instance that:[8]

$$\begin{aligned} U^A_{z_1,2}(y_2,1) &< U^A_{z_1,2}(\hat{y}_2,1) \\ U^A_{z_1,2}(\hat{y}_2,2) &< U^A_{z_1,2}(y_2,2) \end{aligned}$$

Now consider the first period acts $y_1 = (z_1, f_1)$ and $\hat{y}_1 = (z_1, \hat{f}_1)$ where $f_1(s) = y_2$, $\hat{f}_1(s) = \hat{y}_2$ and $f_1(s') = \hat{f}_1(s') = \widetilde{y}^2_2$ for all $s' \neq s$, with $\widetilde{y}^2_2$ defined as in Definition 11. By time consistency, we have $W^A_1(y_1) \geq W^A_1(\hat{y}_1)$. By time consistency, it must also be the case that $U^A_1(y_1,1) < U^A_1(\hat{y}_1,1)$ and $U^A_1(\hat{y}_1,2) < U^A_1(y_1,2)$. And, since s is socially revertible for SO A, we have $U^A_1(\hat{y}_1,2) > U^A_1(\hat{y}_1,1)$.

We end up with $W^A_1(y_1) \geq W^A_1(\hat{y}_1)$ and $y_1 \hat{I}_{U^A_1}\ \hat{y}_1$ because

$$U^A_1(y_1,1) < U^A_1(\hat{y}_1,1) < U^A_1(\hat{y}_1,2) < U^A_1(y_1,2)$$

[8]The case $U^A_{z_1,2}(y_2,1) > U^A_{z_1,2}(\hat{y}_2,1)$ and $U^A_{z_1,2}(\hat{y}_2,2) > U^A_{z_1,2}(y_2,2)$ can be treated similarly using $\widetilde{y}^1_2$ instead of $\widetilde{y}^2_2$ in the remainder of the proof.

Comparative inequality aversion implies that $W_1^B(y_1) \geq W_1^B(\hat{y}_1)$. But given the definition of y_1 and $\hat{y}_1$ and since the SO B is time consistent, this is possible only if $W_{z_1,2}^B(y_2) \geq W_{z_1,2}^B(\hat{y}_2)$.

Then, we have the following lemma:

Lemma 1 *If two SOs have time consistent preferences then:*

If they have the same preferences in period 2 for any z_1, then they have the same preferences in period $t \geq 2$ for any history $h_t \in \mathcal{H}_t$.

Proof. Assume that $W_{h_t,t}^A(y_t) \geq W_{h_t,t}^A(\hat{y}_t)$ with $h_t = (z_1, z_2, \cdots, z_{t-1})$. We need to show that $W_{h_t,t}^B(y_t) \geq W_{h_t,t}^B(\hat{y}_t)$.

Consider the subsequent periods acts defined inductively in the following way: $y_2 = (z_2, f_2)$ with $f_2 \in \mathcal{F}(Y_3)$ and such that $f_2(s) = y_3$ for all $s \in \Sigma$; and, $\forall \tau \in \{4, \cdots, t\}$, $y_{\tau-1} = (z_{\tau-1}, f_{\tau-1})$ with $f_{\tau-1} \in \mathcal{F}(Y_\tau)$ and such that $f_{\tau-1}(s) = y_\tau$ for all $s \in \Sigma$.

Similarly, define the subsequent periods acts in the following way: $\hat{y}_2 = (z_2, \hat{f}_2)$ with $\hat{f}_2 \in \mathcal{F}(Y_3)$ and such that $\hat{f}_2(s) = \hat{y}_3$ for all $s \in \Sigma$; and, $\forall \tau \in \{4, \cdots, t\}$, $\hat{y}_{\tau-1} = (z_{\tau-1}, \hat{f}_{\tau-1})$ with $\hat{f}_{\tau-1} \in \mathcal{F}(Y_\tau)$ and such that $\hat{f}_{\tau-1}(s) = \hat{y}_\tau$ for all $s \in \Sigma$.

By time consistency, $W_{h_t,t}^A(y_t) \geq W_{h_t,t}^A(\hat{y}_t) \Longrightarrow W_{z_1,2}^A(y_2) \geq W_{z_1,2}^A(\hat{y}_2)$. Since A and B have the same preferences in period 2, it is also the case that $W_{z_1,2}^B(y_2) \geq W_{z_1,2}^B(\hat{y}_2)$. The time consistency of B's preferences hence yields $W_{z_1,2}^B(y_2) \geq W_{z_1,2}^B(\hat{y}_2) \Longrightarrow W_{h_t,t}^B(y_t) \geq W_{h_t,t}^B(\hat{y}_t)$. ■

Obviously, Lemma 1 ends the proof. ■

Proposition 2 proves that two time consistent and Paretian SOs with comparable degrees of inequality aversion must have the same preferences after the initial period.

Hence, after the initial period, inequality aversion is constrained. If we consider that period 1 has already occurred, the result is the same as in Proposition 1, but it relies one much weaker grounds, namely time consistency.

A few points deserve comment. First, we have assumed throughout the section that $S > 1$. Otherwise, Assumption 2 cannot be stated. The fact that $S > 1$ is actually crucial to obtain Proposition 2. It is perfectly possible to obtain more inequality averse Paretian SOs in a deterministic intertemporal framework (see Atkinson and Bourguignon, 1982).

A second point is that continuity is not necessary to obtain Proposition 2, contrary to what was the case for Proposition 1. We do not need to make assumptions on the structure of Z_t. We make no assumptions on preferences except representability and time consistency. Hence, Proposition 2 relies on rather weak grounds.

A third point is that state independence is also not necessary to obtain the result. We have assumed state independence for convenience. But it is clear from the proof of Proposition 2 that the result could be obtained in a state dependent framework provided that Assumption 2 is true for any $s \in \Sigma$. More precisely, if all states of the world are sufficiently unimportant so that guaranteeing given allocations in all other states allows individuals' welfare ranks to be inverted, the proof in Proposition 2 ensures that the two SOs have the same (state dependent) ordering in second period.

It is however crucial to assume that unrealized prospects do not matter in second period. Without this consequentialist assumption, the result exposed in Proposition 2 cannot be obtained.

4 Conclusion

We have shown that the time consistency of choices in uncertain consequentialist frameworks implies that we cannot find two social Paretian social observers, with one being more inequality averse than the other one. Two conditions are necessary to obtain this impossibility result: there must be some uncertainty, and we must be able to invert people's welfare ranks in some social alternatives.

We first derived the result in the case of preferences that are linear in mixture. But, we have also shown that even when mixture operations do not exist, the inequality aversion of Paretian SOs is constrained. If we accept the rationality requirements of time consistency and consequentialism, we are thus left with a conflict between individualistic principles and the willingness to reduce inequalities.

The reason is that the Pareto principle forces the social observer to respect individuals' intertemporal choices. But the preferences of individuals in uncertain dynamic frameworks directly impact the inequality of social outcomes. If we endorse the (ex-ante) Pareto principle, the inequalities arising ex-post from people's choices must be accepted, how broad they may be, and there is no room for expressing personal views on the acceptable degree of inequality.

This result leaves us with a difficult dilemma. Either we follow a purely individualistic path and forget about any considerations related to inequality aversion. Or we take into account more redistributive views and have to intervene in individuals' intertemporal decisions. This second path differs from the usual arguments in favor of paternalistic interventions. The Pareto principle is often deemed unappealing when individuals make 'inappropriate' intertemporal decision, for instance when they make dynamically inconsistent choices, or when they have incorrect beliefs, etc. Here, paternalistic interference arises from the ethically defendable concern for equality.

References

Anscombe, F.J., R.J. Aumann, "A definition of subjective probability," *Annals of Mathematical Statistics* 34 (1963), 199-205.

Arrow, K.J., *Aspects of the Theory of Risk-Bearing* (Helsinki: Yrjö Jahnssonin Säätiö, 1965).

Atkinson, A., "On the measurement of inequality," *Journal of Economic Theory* 2 (1970), 244-263.

Atkinson, A., F. Bourguignon, "The comparison of multi-dimensioned distributions of economic status," *Review of Economic Studies* 49 (1982), 183-201.

Bosmans, K, "Comparing degrees of inequality aversion," *Social Choice and Welfare* 29 (2007), 405-428.

Chew, S.H., L.G. Epstein, "The structure of preferences and attitudes towards the timing of the resolution of uncertainty," *International Economic Review* 30 (1989), 103-117.

Diamond, P.A., "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility: Comment," *Journal of Political Economics* 75 (1967), 765-766.

Donaldson, D., J.A. Weymark, "A single-parameter generalization of the Gini indices of inequality,", *Journal of Economic Theory* 22 (1980), 67-86.

Epstein, L.G., U. Segal, "Quadratic social welfare." *Journal of Political Economy* 100 (1992), 691-712.

Fleurbaey, M., "Assessing risky social situations," mimeo, University Paris 5, 2007.

Ghirardato, P., F. Maccheroni, M. Marinacci, "Differentiating ambiguity and ambiguity attitude," *Journal of Economic Theory* 118 (2004), 133-173.

Gilboa, I., D. Schmeidler, "Maxmin expected utility with a non-unique prior," *Journal of Mathematical Economics* 18 (1989), 141-153.

Grant, S., A. Kajii, B. Polak, Z. Safra, "Generalized utilitarianism and Harsanyi's impartial observer theorem," Discussion Paper No. 1578, Cowles Foundation, September 2006.

Grant, S., J. Quiggin, "Increasing uncertainty: a definition," *Mathematical Social Sciences* 49 (2005), 117-141.

Hammond, P.J., "Equity, Arrow's condition, and Rawls' difference principle," *Econometrica* 44 (1976), 793-804.

Hammond, P.J., "Consequentialist foundations for expected utility," *Theory and Decision* 25 (1988), 25-78.

Harsanyi, J.C., "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility," *Journal of Political Economics* 63 (1955), 309-321.

Jewitt, I., "Choosing between risky prospects: the characterization of comparative statics results, and location independent risk," *Management Science* 35 (1989), 60-70.

Johnsen, T.H., J.B. Donaldson, "The structure of intertemporal preferences under uncertainty and time consistent plans," *Econometrica* 53 (1985), 1451-1458.

Karni, E., D. Schmeidler, "Atemporal dynamic consistency and expected utility theory," *Journal of Economic Theory* 54 (1991), 401-408.

Kihlstrom, R.E., L.J. Mirman, "Risk aversion with many commodities," *Journal of Economic Theory* 8 (1974), 361-388.

Kolm, S.-C, “Unequal inequalities I,” *Journal of Economic Theory* 12 (1976), 416-442.

Kooopmans, T.C., “Representation of preference orderings with independent components of consumption,” in T. McGuire and R. Radner, eds., *Decision and Organization* (Amsterdam: North Holland , 1972), 57-78.

Kreps, D.M., E.L. Porteus, “Temporal resolution of uncertainty and dynamic choice theory,” *Econometrica* 46 (1978), 185-200.

Maasoumi, E., “The measurement and decomposition of multi-dimensonal inequality,” *Econometrica* 54 (1986), 991-998.

Machina, M.J., “Dynamic consistency and non-expected utility models of choice under uncertainty,” *Journal of Economic Literature* 27 (1989), 1622-1688.

Myerson, R.B., “Utilitarianism, egalitarianism, and the timing effect in social choice problems,” *Econometrica* 49 (1981), 883-897.

Pratt, J.W., “Risk aversion in the small and in the large,” *Econometrica* 32 (1964), 122-36.

Sarin, R., P. Wakker, “Folding back in decision tree analysis,” *Management Science* 40 (1994), 625-628.

Schmeidler, D., “Subjective probability and expected utility without additivity,” *Econometrica* 57 (1989), 571-587.

Yaari, M.E., “Some remarks on measures of risk aversion and on their use,” *Journal of Economic Theory* 1 (1969), 315-329.