Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 169 (2017) 666–706

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Fair management of social risk ☆

Marc Fleurbaey [a], Stéphane Zuber [b,*]

[a] *Woodrow Wilson School and University Center for Human Values, Princeton University, Princeton, NJ 08544, USA*
[b] *Paris School of Economics – CNRS, Centre d'Economie de la Sorbonne, Maison des Sciences Economiques, 106-112 Boulevard de l'Hôpital, 75647 Paris Cedex 13, France*

Received 8 September 2015; final version received 18 March 2017; accepted 27 March 2017

Abstract

We provide a general method for extending social preferences defined for riskless economic environments to the context of risk and uncertainty. We apply the method to the problems of managing unemployment allowances (in the context of macroeconomic fluctuations) and catastrophic risks (in the context of climate change). The method guarantees ex post fairness and pays attention to individuals' risk attitudes, while ensuring rationality properties for social preferences, revisiting basic ideas from Harsanyi's celebrated aggregation theorem (Harsanyi, 1955). The social preferences that we obtain do not always take the form of an expected utility criterion, but they always satisfy statewise dominance. When we require social preferences to be expected utilities, we obtain a variant of Harsanyi's result under a weak version of the Pareto principle, and a maximin criterion under a stronger Pareto requirement, whenever the ex post social ordering does not depend on people's risk attitudes. We also show how non-expected utility individual preferences can be accommodated in the approach.

JEL classification: D63; D71; D81

Keywords: Social choice; Fairness; Uncertainty; Social risk

☆ This work has been supported by the Agence Nationale de la Recherche (grant numbers ANR-12-INEG-0006-01 and ANR-10-LABX-93); the Chair on Welfare Economics and Social Justice at the Institute for Global Studies (FMSH – Paris); the Franco-Swedish Program on Economics and Philosophy (FMSH & Riksbankens Jubileumsfond); the Princeton Institute for International and Regional Studies.
* Corresponding author.
E-mail addresses: mfleurba@princeton.edu (M. Fleurbaey), stephane.zuber@univ-paris1.fr (S. Zuber).

http://dx.doi.org/10.1016/j.jet.2017.03.009
0022-0531/

1. Introduction

Public policies are fraught with uncertainty about their consequences. To deal with the evaluation of such uncertain consequences, the literature either adopts the expected utilitarian criterion or incorporates risk and uncertainty as a simple extension of the commodity space – commodities being relabeled as contingent on the realization of a state of nature, as proposed in the framework of general equilibrium theory by Debreu (1959, chap. 7). But either approach is questionable.

The expected utilitarian criterion was famously defended by Harsanyi (1955), who proved the following theorem: if (i) all individuals and the social observer are expected utility maximizers on the set of simple lotteries over a set X; (ii) whenever all individuals prefer lottery p to lottery q, the social observer also prefers p to q (the Pareto principle); then the social observer's utility function must be an affine combination of individuals' von Neumann–Morgenstern (VNM) utility functions. This result is known as Harsanyi's aggregation theorem.[1] Although Harsanyi viewed this theorem as a positive result justifying utilitarianism, it is more accurate to understand it as pointing to the tension between rationality (expected utility), *ex ante* Pareto, and equity.

Regarding equity in resource allocation, Harsanyi's theorem is clearly problematic. Consider a one-commodity economy with two individuals 1 and 2 who face a risk on their incomes described by random variables $\tilde{x}_1$ and $\tilde{x}_2$. Assume furthermore that the individuals are expected utility maximizers so that their ordinal preferences are represented by $Eu_1(\tilde{x}_1)$ and $Eu_2(\tilde{x}_2)$. Harsanyi's theorem tells us that society should aim at maximizing $E(\alpha u_1(\tilde{x}_1) + u_2(\tilde{x}_2))$, for some relative weight α. Thus, on riskless prospects, the social welfare criterion should be $\alpha u_1(x_1) + u_2(x_2)$. If the individuals have different risk preferences, this social criterion does not satisfy the Pigou–Dalton transfer principle in general.[2] If u_1 and u_2 are the same and $\alpha = 1$, then the Pigou–Dalton transfer principle is satisfied but inequality aversion is constrained by risk aversion, which does not seem particularly appealing either.

Starting from a criterion of equity in resource allocation, one could think that it is natural to adopt the other approach invoked in the first paragraph, dealing with contingent commodities. The problem is that such an approach is likely to entail a violation of rationality at the social level. Indeed, an act that increases inequality ex post, thereby making sure that the final consequence is worse in every state of the world, may appear attractive ex ante if everyone has a reasonable chance of being on the winning side. A typical fair criterion extended to acts in this way will therefore often violate the basic rationality condition of statewise dominance. This violation appears to be a direct consequence of the combination of the ex ante Pareto principle with fairness considerations, as illustrated in the extreme case of an egalitarian society by Gajdos and Tallon (2002). The tension between rationality, ex ante Pareto and fairness is therefore serious, and one must seek compromises that trade off these important values.

In this paper, we examine how to evaluate risky prospects while preserving basic rationality conditions (either statewise dominance or expected utility) as well as basic fairness principles. Our approach is compatible with a wide range of social criteria for the evaluation of non-risky allocations, involving various types of interpersonal comparisons and degrees of inequality aversion. In particular, we are not committed to the assumption that individuals' expected utilities provide the correct measure for assessing and comparing individual welfare.

[1] Several versions of the theorem have been established for different expected utility models in frameworks involving risk or uncertainty (Blackorby et al., 1999; Fishburn, 1984; Mongin, 1995, 1998).

[2] At any value x such that $\alpha u_1'(x) \neq u_2'(x)$, a small inequality in favor of the individual with greater marginal utility is better than equality.