

# Stéphane Zuber, Équité, Risque et Temps : Rendre Justice au Futur (HDR)

Je résume ici le mémoire d'habilitation à diriger des recherches de Stéphane Zuber (2017). Je n'ai pas reproduit les équations, ce qui peut compliquer la compréhension du texte. Dans tous les cas, je ne pense pas que ce résumé soit très compréhensible pour quelqu'un qui n'a pas lu les travaux ou l'HDR de S. Zuber. Ces travaux portent sur la formalisation de l'éthique. Plus précisément, ils montrent que pour intégrer le risque de façon satisfaisante dans une fonction de bien-être social, on doit renoncer au conséquentialisme (en tenant compte des états du monde non réalisés), et que la version forte du principe de Pareto n'est pas compatible avec le principe d'équité. En utilisant des fonctions de bien-être dépendant du rang pour traiter le cas inter-temporel, on peut dépasser cette incompatibilité. Ces types de fonctions permettent d'éviter les implications problématiques du prioritarisme dans le cadre d'une population variable. Enfin, l'auteur généralise l'actualisation grâce à un taux d'escompte social dépendant de l'utilité marginale des contributeurs et des bénéficiaires d'une politique.

## 0. Contexte

On s'intéresse à des agrégateurs de bien-être  $F_w$ , dans le cadre welfariste où on suppose les niveaux de bien-être individuel comparables et cardinaux.

Deux approches seront employées : l'approche prioritariste, où on somme une fonction concave des bien-être ; et les fonctions de bien-être dépendant du rang, où on pondère les utilités individuelles par un coefficient dépendant de leur rang.

## 1. Risque et Équité

En présence de risque, Harsanyi (1955) montre que, dès lors qu'on respecte le principe de Pareto et qu'on définit le bien-être social comme l'utilité espérée, on peut écrire la fonction de bien-être social comme l'espérance de la somme (d'une transformation affine) des utilités individuelles.. Ça pose problème (Diamond 1967, Broome 1991) car on somme indistinctement selon les états du monde et les individus, entraînant une indifférence face aux inégalités. Pour satisfaire le principe d'équité, l'auteur étudie des affaiblissements du principe de Pareto.  $f$  est une « allocation incertaine », i.e. une variable aléatoire assignant une allocation, fonction des états du monde. Les individus et le planificateur ont une fonction d'utilité VNM s'il existe une fonction  $U$  telle que l'espérance de  $U(f)$  définit une relation de bien-être social (qui classe les allocations incertaines). Un principe de Pareto affaibli est que  $f$  est préférée à  $g$  (noté  $f R g$ ) ssi pour tout individu  $i$  les espérances de  $u_i(f) > u_i(g)$  pour toutes allocations  $f$  et  $g$  égalitaires ou certaines (i.e. ne dépendant pas de l'état du monde).

Le principe de transfert (ici, multidimensionnel) dit que, toute chose égale par ailleurs, si dans tous les états du monde une allocation est préférée lorsqu'elle rapproche de la même variation d'utilité l'utilité de deux individus. Fleurbaey (2010) a montré que, lorsque la relation de bien-être social est représentée par une espérance d'utilité, la combinaison de ces deux principes conduit à évaluer les allocations incertaines en fonction de la valeur espérée d'un niveau d'utilité équivalent et distribué de façon égalitaire. Le problème est qu'une telle fonction de bien-être social est fortement non-séparable entre sous-populations : l'allocation équivalente distribuée de façon égalitaire dans un état de la nature dépend généralement de l'ensemble de la distribution du bien-être dans cet état (exemple : maximin). L'indépendance de l'utilité des morts est souhaitable, et correspond cependant à une propriété plus faible de séparabilité que la séparabilité entre sous-populations mentionnées plus haut. En effet, les morts ne font plus face à aucun risque. Le principe d'indépendance de l'utilité des personnes ne faisant pas face au risque stipule que  $f R g$  ssi  $F R G$ , dès lors qu'il existe une sous-population  $M$  (ne faisant

pas face au risque) telle que les allocations de même casse coïncident sur  $M$ , et que les allocations de même lettre coïncident sur le complémentaire de  $M$  (en gros, on peut séparer les morts du reste de la population et négliger leur bien-être). Zuber présente alors une alternative au théorème d'Harsanyi, lorsque les trois principes ci-dessus sont respectés : les utilités sont alors multipliées plutôt que sommées, sous l'espérance. Il en découle que la préférence pour la redistribution peut exister mais de façon limitée, loin du cas égalitariste. Plus problématique, le bien-être social dépend de la taille de la population (morte), et n'est donc pas stationnaire, i.e. indépendant de l'existence des morts. Si l'on veut une forme assez forte d'indépendance par rapport au passé, on est contraint de revenir à un critère utilitariste qui est indifférent vis-à-vis de la redistribution du bien-être entre les individus.

Ensuite, Zuber introduit un principe d'équité *ex ante*, la préférence pour la randomisation, qui dit qu'une allocation incertaine est préférée à une allocation certaine, si elle consiste en une randomisation des individus de l'allocation certaine. Ce principe impose de renoncer au conséquentialisme (et à la parcimonie de l'information requise).

Je passe les quelques pages où l'auteur montre comment se ramener à des allocations certaines et égalitaires pour généraliser la théorie aux fonctions d'utilité non VNM.

## II. Équité entre générations et éthique de la population

Diamond (1965) montre que le principe de Pareto est incompatible avec le principe d'égal traitement de toutes les générations, dans le cadre de vecteurs infinis de bien-être. Mais l'utilité escomptée fait sens, si chaque génération est mieux lotie que la précédente : l'actualisation s'interprète alors comme une pondération. Zuber & Asheim (2012) définissent les relations utilitaristes escomptées par le rang comme l'utilitarisme escompté après ré-ordonnement des utilités de sorte qu'elles croissent dans le temps ; cette relation n'est définie que sur les allocations inter-temporelles pouvant être réordonnées. Ces relations satisfont le principe de Pareto fort ainsi qu'une version forte du traitement égal des générations, dépassant les impossibilités mises à jour par la littérature. En affaiblissant à peine le principe de Pareto, on peut étendre le domaine de ces relations et leurs propriétés à l'ensemble des allocations, et d'autres propriétés sont mêmes vérifiées (séparabilité, stationnarité restreinte).

Lorsqu'on intègre la finitude et la variabilité de la population, toutes les approches prioritaristes échouent à nous satisfaire (ex : conclusion répugnante pour l'utilitarisme classique ; la violation de la priorité aux vies valant d'être vécues pour l'utilitarisme avec niveau critique  $c > 0$ , car l'ajout d'un individu d'utilité  $0 < u < c$  n'est pas souhaitable). Utiliser des fonctions dépendant du rang permet de surmonter ces difficultés, cf. Asheim & Zuber (2014). Les relations utilitaristes escomptées par le rang garantissent qu'il existe des vecteurs égalitaires aussi grand que l'on souhaite et équivalents à une population finie avec un niveau de bien-être donné. Cette propriété est essentielle pour éviter la conclusion répugnante: la taille de population a une contribution finie au bien-être social. On obtient une invariance à la réplification des populations avec une pondération de type  $p_r^d - p_{r-1}^d$  où  $p_r$  est la probabilité d'existence du  $r^{\text{ème}}$  individu le moins bien loti (en terme de l'ordre leximin utilité, proba). Contrairement au prioritarisme avec niveau critique et pondération de la population, les relations de bien-être dépendant du rang peuvent à la fois respecter le principe d'expansion négative (il n'est pas souhaitable d'ajouter un individu à utilité négative), la priorité aux vies valant d'être vécues (on préfère un groupe d'individus à utilités positives qu'un groupe à utilités négatives) et le principe fort d'addition de qualité (on préfère ajouter un certain nombre d'individus mieux lotis qu'un nombre arbitrairement grand d'individus moins bien lotis). Mais les relations dépendant du rang ne satisfont pas le principe d'addition (on préfère ajouter un individu à utilité positive), car il respecte le principe (égalitariste) faible de transfert, et qu'on ne veut pas ajouter un individu très pauvre, qui exacerberait l'inégalité. Les auteurs étendent ces notions aux situations où l'existence est probable et non plus certaine, et montre qu'une relation satisfait un certain nombre de propriétés ssi  $c$  est une relation utilitariste escomptée par le rang et ajustée pour les probabilités d'existence (avec un facteur similaire à  $p_r^d - p_{r-1}^d$ , et un niveau

critique). Ces propriétés sont le principe de Suppes-Sen, qui combine l'anonymat et Pareto, l'indépendance de l'existence des mieux et des moins bien lotis (les intégrer dans deux allocations ne change pas la préférence entre celles-ci), l'existence faible d'un niveau critique, et le principe de transfert.

### III. Le taux d'escompte social

Fleurbaey & Zuber introduisent un taux d'escompte inter-personnel entre l'individu  $i$  de la génération 0 et l'individu  $j$  de la génération  $t$ , qui permet de fonder une notion plus générale de taux d'escompte social permettant de rendre comparable la valeur de la consommation (agrégée) actuelle et celle d'impacts futurs attendus d'une politique. Le taux d'escompte social dépend alors des individus concernés par la politique, et ne peut s'exprimer comme une simple moyenne généralisée des taux d'escompte inter-personnels, car il doit tenir compte de la corrélation entre les facteurs d'escompte inter-personnels définis état du monde par état du monde et les parts des rendements (et non des niveaux car le rendement total est supposé certain) allant aux individus. Les auteurs ont étudié plus en détail ces taux d'escompte social pour une large de classe de fonctions de bien-être équivalentes prioritaristes, définies comme une combinaison linéaire (où les coefficients dépendent de la taille de la population) d'un niveau d'utilité équivalent distribué de façon égalitaire, qui s'écrit comme l'utilité moyenne transformée par une fonction concave sous la somme et par son inverse en dehors de la moyenne. Cette fonction concave permet de représenter l'aversion pour l'inégalité, tandis que l'utilité individuelle représente l'aversion pour le risque. Les auteurs présentent alors une formule pour le taux d'escompte inter-personnel dans le cadre de ces fonctions de bien-être. Cette formule complexe est grandement simplifiée sur le très long terme, car elle converge vers le taux d'escompte inter-personnel entre le plus riche vivant et le plus pauvre du futur. Fleurbaey & Zuber (2015) bornent la vitesse de convergence dans le cadre d'investissements certains. Ils montrent par un exemple simple que la convergence est atteinte au bout de 1000 ans, et que la marge d'erreur est inférieure à 1 % à tout horizon. Dans d'autres papiers, l'auteur étudie les situations qu'on ne peut pas représenter par des effets marginaux et où on ne peut donc pas employer un taux d'escompte, et montre comment utiliser une valeur sociale de la réduction du risque catastrophique, qui est le pendant pour les sociétés de la valeur statistique d'une vie pour les individus. Tous ces développements théoriques plaident pour une plus forte réduction des émissions de gaz à effet de serre.

#### Bibliographie :

Arrhenius, à paraître  
Blackorby et al., 2005  
Dasgupta-Heal-Solow  
Gollier, 2015  
Harsanyi, 1955  
Savage, 1954