

L'économie contemporaine dispose-t-elle d'une conception satisfaisante de l'utilité?

Conférence préparée pour l'Université d'été: "Utilité et bien-être: histoire, méthodes, mesures", Strasbourg, Pôle Européen de Gestion et d'Economie, 1-5 septembre 2003.

Philippe MONGIN

C.N.R.S et Laboratoire d'économétrie, Ecole Polytechnique

La question posée dans le titre en contient plusieurs, auxquelles on peut répondre en séquence:

- L'économie contemporaine dispose-t-elle véritablement d'une conception de l'utilité?
- A supposer que cette conception existe, est-elle cohérente avec les objectifs théoriques de l'économie contemporaine?
- En supposant que ce ne soit pas le cas, l'économie contemporaine pourrait-elle rétablir sa cohérence interne sur une autre base conceptuelle?

A la première question, l'exposé répondra positivement; à la seconde, négativement; à la troisième positivement. L'argument négatif, le plus développé, consiste à démontrer que *l'économie contemporaine se met en conflit avec elle-même, en développant et valorisant des recherches qui ne peuvent pas tolérer sa conception de l'utilité*. Les contre-exemples concernent (i) l'aversion pour le risque et la probabilité subjective en théorie de l'utilité espérée, et (ii) l'optimalité parétienne dans la théorie du bien-être économique.

La conclusion finale, qui n'est pas pessimiste, recommande un abandon partiel de la conception reçue. Elle en préserve certaines thèses et ne fait pas passer au second plan l'utilité, contrairement aux développements "post-welfaristes" d'aujourd'hui. L'effet immédiat serait d'intégrer des recherches importantes qui sont reléguées dans les marges.

Avertissements préliminaires.

- On ne mettra pas au centre l'interprétation de la "préférence révélée". Si elle était simplement identique à la conception contemporaine, la réfutation serait trop facile!

- Un développement qui renforcerait les conclusions négative et positive est purement et simplement omis (sur l'utilité cardinale).

-Pour éviter une circularité des deux premières questions, on prend la première comme une *question de fait*, et il faut alors délimiter approximativement le corpus:

- des manuels ou traités d'aujourd'hui qui font autorité en micro-économie, théorie de la décision, économie du bien-être (Malinvaud, Mas-Colell, Winston, Green=MWG, Fishburn, Boadway & Bruce);
- les travaux des fondateurs (Hicks, Bergson et la NEBE, Samuelson, von Neumann & Morgenstern, les théoriciens de l'utilité VNM, Savage, Luce, Suppes et collaborateurs en théorie du "mesurage", Arrow, Chipman, Hurwicz et d'autres réunis dans *Preferences, Utility, and Demand*, 1971).

L'exposé ne retrace pas la genèse de la conception contemporaine, cependant une interprétation historique (simplement affirmée ici) l'accompagne:

- La microéconomie positive reste principalement néo-parétienne, ce qui peut convenir à la théorie du consommateur, telle justement que les néo-parétiens l'ont développée, mais qui crée des tensions sérieuses au-delà, p.ex. en théorie du risque.
- La microéconomie normative reste sous la dépendance intellectuelle de la "nouvelle économie du bien-être" (NEB) et de son projet réducteur, même si cette école est officiellement discréditée.
- La théorie de la décision, créée dans l'après-guerre, n'a pas joué tout le rôle clarificateur attendu, à cause de la condescendance qu'elle a montrée à l'égard des questions normatives et de positions béhavioristes qu'elle a prises en théorie positive.

1 Diagnostic de la conception contemporaine (CC) de l'utilité

1. Pour la CC, l'utilité est un concept non pas premier, mais dérivé. Le concept dont elle dérive est la préférence. La dérivation s'entend d'abord *au sens logico-mathématique des théorèmes de représentation*, tels que la théorie de la décision en a fixé le modèle. La fonction d'utilité est un objet numérique; la préférence est une relation, définie avec une aide minimale des nombres ("qualitativement").
2. L'utilité est en outre un concept *sémantiquement* dérivé. Le terme primitif "relation de préférence" est directement interprétable, et ses interprétations procurent au terme "fonction d'utilité" *toute* sa signification. L'utilité s'interprète alors comme le degré de satisfaction des préférences.
3. Il en résulte que les propriétés éventuelles de la fonction d'utilité - linéarité, concavité, etc, ainsi que l'unicité relative par rapport à telle ou telle propriété - ne prennent de signification que si l'on retourne à l'objet plus fondamental qu'est la préférence.

4. La préférence est liée conceptuellement au choix, sans que les deux concepts de préférence et de choix puissent être considérés comme équivalents. La formule qui les lie au sein de la CC n'est pas explicite. Ce qu'on appelle "théorie de la préférence révélée" introduit des confusions qu'il faudra dissiper.
5. Au travers de la préférence, la fonction d'utilité tire des choix le moyen de sa *détermination empirique* (on peut dire aussi: de sa *mesure*). La CC admet que le choix est l'objet qui se prête le plus à l'observation, mais il peut y avoir des restrictions spéciales à imposer pour que des choix passent pour observables.
6. En théorie normative, la préférence et l'utilité sont associées à un concept de *désirabilité subjective*, individuelle ou collective, que la CC appelle "bien-être". La nature sémantique ou non de l'association n'est pas claire, mais la CC tient pour acquis qu'elle vient se rajouter, et non se substituer, aux considérations précédentes. C'est la préférence et l'utilité *satisfaisant aux conditions 1-5* qui interviennent en théorie normative.

On n'inclura pas dans la CC une position sur la *cardinalité* de l'utilité individuelle, ni une position sur la *comparabilité* des utilités d'individu à individu. Les deux mots sont pris au sens technique de la théorie de la décision ou du choix social.

Une fonction d'utilité u est ordinaire si elle est unique à une transformation croissante près, et cardinale si elle est unique à une transformation affine positive (TAP). (D'autres définitions mathématiques de la cardinalité existent, celle-ci étant un peu trop forte, mais c'est la plus fréquente.)

Deux fonctions u et v sont comparables ordinalement si elles sont uniques à une transformation croissante commune près, comparables cardinalement si elles sont uniques à une TAP commune près, etc.

Les économistes ont besoin de cette notion de cardinalité pour la théorie du risque et de la probabilité subjective, ainsi que pour certaines applications d'économie normative (les métriques monétaires). Ils ont besoin de cette notion de comparabilité pour que la fonction de bien-être social échappe à la trivialité (cf. la "réfutation" par Arrow de la fonction de Bergson-Samuelson).

Une clarification doit porter sur la “**théorie des préférences révélées**”.

L'expression peut désigner une **démarche logico-mathématique**. En théorie du consommateur, elle consiste à imposer à la demande l'axiome fort de la préférence révélée, d'où l'on déduit les propriétés classiques de la demande contenue dans les relations de Slutsky. Les manuels signalent cette dérivation, due à Samuelson et Houthakker, mais à titre subordonné, et de telles dérivations *des choix vers les choix* sont de toute façon rares en dehors de la théorie du consommateur. Une autre démarche logico-mathématique, liée à la précédente, consiste à partir ou bien de l'axiome fort, ou bien des relations de Slutsky, pour prouver l'existence d'une fonction d'utilité ordinale, en appliquant un théorème d'intégrabilité. Dans ce cas, on dispose d'une dérivation de l'utilité à partir des choix directement, sans mention de la préférence.

“Théorie des préférences révélées” désigne souvent non pas une démarche logico-mathématique, mais une **sémantique** de la “préférence” tendant à rendre ce terme synonyme de celui de “choix” (ou l'amalgame de la démarche logico-mathématique avec cette sémantique).

La sémantique de la préférence révélée: les préférences n'ont pas de signification indépendante des choix. La carte des préférences de l'agent serait un résumé commode de l'ensemble des choix qu'il est susceptible d'effectuer quand on fait varier son ensemble des possibilités. *La démarche logico-mathématique n'implique évidemment pas la sémantique.*

La dominante néo-parétienne vers 1930-1960 (Hicks) et la CC aujourd'hui (MWG, etc) répondent à une conception de la préférence qui lui attribue *l'existence autonome et même un rôle causal sur les choix*, d'où la différence sémantique des deux. Il reste que la démarche logico-mathématique des préférences révélées a joué un rôle décisif dans les développements du néo-parétianisme. Le rôle de Samuelson doit être réapprécié.

Voici une formule acceptable pour la CC: "Si i est parfaitement informé des objets de choix x et y , que i veuille exclusivement satisfaire ses préférences, et que i ait la possibilité de choisir entre x et y , alors i choisit x plutôt que y si et seulement si i préfère strictement x à y ." La théorie du consommateur se donne justement les conditions réclamées pour i .

2 L'attitude par rapport au risque et la probabilité subjective

La conception relativement souple et vague qui a été décrite est *trop contraignante encore*. Les difficultés commencent avec les points 3), 4) et 5).

Dans la théorie VNM de l'utilité espérée, les axiomes portent sur les préférences exprimées entre des billets de loterie. Ils impliquent l'existence d'une fonction d'utilité définie sur les mêmes objets et ayant la forme de l'utilité espérée $\sum p_j u(x_j)$. Toute autre fonction d'utilité présentant cette forme résulte de la précédente par une TAP:

$$\sum p_j v(x_j) = \sum p_j (a u(x_j) + b).$$

La dernière affirmation ne signifie pas seules les fonctions ayant la forme de l'utilité espérée conviennent pour représenter la préférence! La forme générale est $\varphi(\sum p_j u(x_j))$ pour n'importe quelle φ croissante, affine ou non.

Quand les résultats x sont monétaires, la concavité de u prend la signification de l'aversion pour le risque et le coefficient d'Arrow-Pratt, $A(x) = -\frac{u''(x)}{u'(x)}$, qui est une mesure numérique de la concavité, sert aussi de mesure à l'aversion pour le risque. La justification repose sur des théorèmes qui relient la décroissance de $A(x)$ à une plus grande disposition à prendre des risques, *cette disposition étant décrite en termes de préférences seulement*. L'agent qui a de l'aversion pour le risque en x préfère à toute loterie autour de x un montant certain inférieur à la valeur actuarielle de cette loterie (prime de risque positive), etc.

$A(x)$ est invariant par rapport aux TAP seulement, et non pas par rapport aux φ quelconques, mais cela suffit pour que le coefficient soit lié à la préférence. (Deux représentations ont la forme $\varphi(\sum p_j u(x_j))$ et $\psi(\sum p_j v(x_j))$, v étant une TAP de u .)

Ces résultats sont conformes à la CC, à condition qu'elle passe sur les hypothèses de *dérivabilité*: celles-ci n'ont pas de signification du côté de la préférence et sont donc littéralement injustifiées en vertu de 3). Le problème se repose dans la théorie de l'utilité espérée généralisée (Machina).

Dans la théorie de Savage, les axiomes portent sur les préférences entre des “actes”, cad des fonctions $f : S$ (états du monde) $\rightarrow X$ (conséquences). Les axiomes impliquent l’existence d’une probabilité π sur les états et d’une fonction d’utilité sur les actes qui a la forme de l’utilité espérée relativement à cette mesure: $\sum \pi_s u(f(s))$. Toute autre représentation ayant la forme mathématique comportera, pour la probabilité, la même π , et pour l’utilité, une TAP de u . L’apport de Savage est qu’il *déduit* une probabilité au lieu de la poser, comme faisaient VNM, et qu’elle est interprétable comme probabilité subjective à cause de son *unicité relative*.

L’axiomatique n’énonce pas de conditions nécessaires pour sa conclusion. Il apparaît la difficulté technique des *conditions non ouvertes*, qui restreignent la classe des ensembles d’options autorisés, par rapport à tous ceux sur lesquels on peut définir une fonction d’utilité ayant la forme requise. Savage a besoin d’un S qui présente une certaine propriété de divisibilité par rapport à la préférence (postulat 6). Toutes les axiomatiques de la probabilité subjective contiennent une condition de “solvabilité”, cf Luce, Suppes et alii. C’est un écart entre les concepts de préférence et d’utilité espérée qui, littéralement, réfute le point 3).

La plus grande difficulté vient d'un raisonnement pourtant trivial: on peut remplacer $\sum \pi_s u(f(s))$ par $\sum \pi'_s v(s, f(s))$ en fixant arbitrairement les $\pi_s > 0$ et en définissant pour tout s :

$$v(s, f(s)) = \frac{p_s u(f(s))}{q_s}.$$

On ne change pas la représentation dans le théorème d'existence, mais le théorème d'unicité disparaît puisqu'il y a une infinité de choix possibles pour π' . On a rendu l'utilité dépendante des états en prenant des $v(s, \cdot)$, mais *rien dans l'axiomatique de Savage n'impose de retenir exclusivement une représentation indépendante des états* $u(\cdot)$.

La notion de probabilité subjective est ébranlée, et avec elle, l'analyse de l'aversion pour le risque quand X représente la monnaie. Si la concavité de $v(\cdot, x)$ signifie l'aversion pour le risque en x , il faut que le détenteur de fonctions $v(\cdot, x)$ concaves préfère à f un montant inférieur à sa valeur actuarielle $\bar{x} = \sum \pi_s f(s)$, mais on ne sait plus avec quelle π calculer cette valeur.

La solution généralement proposée au "problème de l'utilité dépendante des états" consiste à renforcer l'axiomatique de Savage pour restreindre les fonctions $v(s, \cdot)$ et, indirectement, π , mais *cela conduit à introduire des préférences d'un nouveau genre*.

- Chez Karni et Mongin, l'agent doit pouvoir exprimer une préférence entre "recevoir 100 euros alors que la température est de 40°" et "recevoir 50 euros alors que la température est de 25°" (en fait, sur des billets de loterie construits sur ces résultats). Or cette préférence ne correspond pas à un choix possible puisque les deux options ne seront jamais simultanément disponibles. La construction viole la sémantique de la préférence révélée, et ne s'accorde avec celle de la CC (=point 4) que *trivialement* (en niant l'antécédent).
- Chez Drèze, l'agent n'a de préférences que sur les actes, mais il influence les états: l'archer préfère-t-il *f* qui lui donne 1000F s'il atteint la cible ou *g* qui lui donne 50 F s'il la manque? La construction reste compatible avec 4), mais au prix d'une hypothèse restrictive inattendue. *C'est une autre forme de contrefactualité, imposée aux choses mêmes!*

On ne peut pas non plus garder le point 5) sur la mesure. Les choix qui permettraient de mesurer la probabilité subjective et, finalement, l'attitude par rapport au risque ne sont pas du type observable. Il faut accepter de construire des mesures *à partir de données verbales et d'expériences de pensée.*

3 L'optimalité parétienne et le principe de Pareto (PP)

En matière normative, on va montrer que la CC ne parvient pas même à défendre sa conception la plus ordinaire, celle de l'optimalité parétienne! L'échec est d'autant plus grave qu'il ne survient pas à l'occasion d'une extension inattendue de la théorie. La cause en est encore la surdétermination des conditions, ici liée à l'adjonction de 6).

L'optimum et le principe de Pareto se définissent à l'aide du concept de préférence:

(i) Un optimum de Pareto est un état réalisable de l'économie x tel qu'il n'existe pas d'autre état réalisable y qu'un membre de la société au moins préfère strictement à x et que les autres préfèrent faiblement à x .

(ii) Si un membre de la société au moins préfère strictement x à y et que les autres préfèrent faiblement x à y , alors la société doit préférer x à y .

(On passera sur les variantes faibles de (i) et (ii), et sur la version de (ii) en termes d'indifférence.)

Les affirmations (i) et (ii) portent sur les préférences *naturelles* des individus, et non pas sur des préférences “réfléchies”, “élaborées”, “améliorées”, “censurées”, comme chez les philosophes analytiques (Griffin) ou dans le néo-utilitarisme (Harsanyi). *La CC est formelle là-dessus (point 6)*. Les économistes normatifs qui s’en éloignent le font sciemment (Mongin & d’Aspremont, Blackorby, Bossert & Donaldson).

Avec cette notion de préférence, *l’optimum* n’est pas encore un concept normatif. Il en résulte que les “deux théorèmes fondamentaux de l’économie du bien-être” peuvent se lire comme des propositions d’économie positive (c’est le meilleur argument, à distinguer d’autres qui ont cours - comparer avec Archibald).

Même avec cette notion de préférence, l’énoncé du *principe* est déjà normatif. La seule manière d’éviter cette conclusion serait de lire “doit” au sens de la nécessité logique et non du devoir, c’est-à-dire de prendre (ii) comme une définition partielle de ce qu’est l’objet: préférence sociale. Cette interprétation analytique ne se rencontre pas dans la CC.

Le caractère normatif de l'optimum résulte d'une reformulation de (i) en termes de bien-être (BE):

(i') Un optimum de Pareto est un état réalisable de l'économie x tel qu'il n'y ait pas d'autre état réalisable y qui, pour un membre de la société au moins, comporte plus de BE et qui, pour les autres, n'en comporte pas moins.

La CC traite (i) et (i') comme équivalents avec un minimum de justification (inexistante chez MWG, une phrase chez Boadway & Bruce, etc). Elle sous-entend que "le BE" n'est pas toujours à prendre au sens ordinaire, mais désigne éventuellement ce qui est désirable, bon, du point de vue de l'individu considéré. Quel qu'en soit le sens, la liaison postulée entre préférences et BE est censée fournir sa valeur normative au concept (i) et sa justification au principe normatif (ii). *C'est un autre aspect de la CC (point 6).*

On vient de rappeler les concepts parétiens en veillant à les rendre compatibles avec la CC (et la NEBE qui l'inspire). On va voir que cette version du parétianisme n'est pas défendable.

Il y a trois manières d'approfondir ce qui précède, chacune impliquant une défense du principe de Pareto.

- La thèse (A) prend la notion de BE au sens ordinaire et la lie *sémantiquement* à la satisfaction des préférences naturelles. Dans la version crue: elles auraient le même sens. Si la thèse était correcte, elle pourrait justifier l'universalité donnée au principe de Pareto. Mais il y a l'objection évidente: les objets qui font le BE d'un individu ne sont pas toujours ceux qu'il préfère naturellement. Seules des préférences "élaborées" par *la certitude parfaite et l'annulation des externalités* peuvent avoir un rapport avec le BE.
- La thèse (B) prend encore la notion de BE au sens ordinaire et *distingue significations et mesures*. Le BE ne veut pas dire la satisfaction des préférences, mais les deux coïncident dans les circonstances convenables pour la théorie économique, de sorte que celle-ci peut prendre l'une comme mesure de l'autre. (Penser à l'analogie de la masse et du poids.) C'est en fait la thèse de l'EBE sophistiquée (Little, Boadway & Bruce).

Discussion de la thèse (B):

(a) Elle ne justifie pas le PP au niveau de généralité où il est posé, alors que la thèse sémantique aurait permis de le faire.

(b) Que sont les “circonstances convenables pour la théorie économique”? Il faut qu’elles coïncident avec celles où les préférences élaborées émergent naturellement, sinon l’objection précédente revient. Or les conditions de certitude parfaite et d’annulation des externalités sont rarement réalisées.

(c) L’EBE peut préserver logiquement, sinon raisonnablement, une interprétation des deux “théorèmes fondamentaux” en termes de BE parce qu’elle fait alors des hypothèses de certitude et d’absence d’externalités. En dehors de cette application, (b) rend l’interprétation en termes de BE inutilisable

- Une troisième thèse, (C), prend la notion de BE dans le sens de ce qui est bon pour l’individu et cherche une justification grâce à l’élargissement du sens. Le PP reposerait sur celui-ci: “Si x est plus conforme que y à ce qui est désirable pour chaque membre de la société, alors la société doit préférer x à y .” Mais le saut est visible, la préférence naturelle portant sur ce qui est désiré, non ce qui est désirable.

D'autres défenses du PP ne font pas intervenir le bien-être du tout. Les unes renversent la formule de l'EBE: il se justifie non parce que les préférences portent sur le bon, mais parce qu'il est bon de respecter les préférences (sous-entendu: quelles qu'elles soient). On peut développer la thèse d'une manière pragmatique (avantages de "concrétisation" par exemple) ou carrément déontologique (on valorise la liberté dans l'exercice de la préférence). L'interprétation *libérale* du PP suit la dernière direction; ex. du débat américain sur les ceintures de sécurité.

Une autre manière de sauver quelque chose du PP est de le fonder sur le bien-être, mais en le reformulant à partir des préférences élaborées. *L'utilitarisme du XX^e siècle prend cette direction quand il est conséquent.* Il est "préférentialiste" tout en dépendant d'une idée de bonheur; seules les préférences élaborées peuvent faire le lien. (Il reste à clarifier la notion de cardinalité convenable pour le néo-utilitarisme, problème que Harsanyi n'a pas su résoudre contrairement au précédent.)

Dans tous les cas la justification du PP fait perdre quelque chose par rapport à la CC: l'interprétation libérale sacrifie le point 6), cad la référence traditionnelle au BE, et l'interprétation néo-utilitariste sacrifie le point 4), cad la sémantique des préférences.

Il y a une raison indépendante de contester 4) quand on aborde l'économie normative: le fait que l'objet typique d'évaluation est *l'état de la société*, et qu'il n'est pas, en tant que tel, objet de choix pour l'individu. Certains diraient qu'il n'est pas même objet de choix, mais simplement d'évaluation, pour la puissance publique ou son représentant. L'objection ne porte pas seulement contre la sémantique de la préférence révélée, que l'EBE écarte prudemment, mais contre la sémantique plus subtile de la CC.

La mesure disparaît en même temps que la sémantique, et avec elle une grande part de l'EBE *appliquée* telle qu'on la pratique, avec les procédés de la disponibilité à payer, du surplus du consommateur, etc. Ces procédés sont de toute façon théoriquement contestés à l'intérieur même de l'EBE.

4 Conclusions

Les points 1-5 imposent une contrainte insupportable à l'économie positive. La théorie néo-parétienne du consommateur pouvait s'accommoder des cinq points, mais elle n'était que *le banc d'essai* de la CC. L'extension de la micro-économie du côté du risque et de l'incertitude force à revoir la sémantique de l'utilité en termes de préférence et, surtout, de la préférence en termes de choix. Les néo-parétiens avaient fait un pas en mettant à distance l'interprétation de la préférence révélée. On suggère aux économistes contemporains d'aller plus loin dans la même direction.

Le point 6 met l'économie normative dans une position intenable. La difficulté, cette fois, n'est pas due à un développement supplémentaire de la théorie initiale, mais à la manière même - incohérente - dont elle s'est constituée. Les néo-parétiens, en l'occurrence les NEBE, portent ici une pleine responsabilité. La recommandation consiste de nouveau à sacrifier la sémantique de la préférence en termes de choix. Il resterait à dire quelles *données supplémentaires* conviennent, dans cette perspective, à l'EBE appliquée.