

ENRICO DIECIDUE

DERIVING HARSANYI'S UTILITARIANISM FROM DE FINETTI'S BOOK-MAKING ARGUMENT

ABSTRACT. The book-making argument was introduced by de Finetti as a principle to prove the existence and uniqueness of subjective probabilities. It has subsequently been accepted as a principle of rationality for decisions under uncertainty. This note shows that the book-making argument has relevant applications to welfare: it gives a new foundation for utilitarianism that is alternative to Harsanyi's, it generalizes foundations based on the theorem of the alternative, and it avoids arguments based on expected utility.

KEY WORDS: book-making argument, policy independence, utilitarianism

JEL CLASSIFICATION NUMBER: D60, D71, D81

1. INTRODUCTION

This note presents a consistency condition for welfare policies. It is based on the book-making argument introduced by de Finetti (1931), a famous argument in decision under uncertainty. The argument states a natural condition that turns out to imply the existence of coherent subjective probabilities and justifies a model of choice based on them. De Finetti's idea served as point of departure for Savage's (1954) theory of subjective expected utility.

This note shows that de Finetti's book-making argument also has relevant applications in welfare, where it provides a natural foundation for utilitarianism that is alternative to Harsanyi's (1955) and has an advantage over Harsanyi's by avoiding reliance on expected utility (von Neumann and Morgenstern, 1947). Harsanyi's approach and its dependence

on expected utility is nowadays criticized, a modification of the approach is desirable and highly called for (Ebert, 1988; Weymark, 1981; Yaari, 1988). Our foundation is desirable in light of the modern work on deviations from expected utility (Kahneman and Tversky, 1979; Schmidt, 2004; Starmer, 2000).

In a mathematical sense, this note generalizes earlier foundations of utilitarianism based on the theorem of the alternative by using natural combinations of policy decisions, and not unnatural scalar multiplications. The main contribution of this note does not concern, however, its mathematical generality but the high desirability of a clear restatement of Harsanyi's famous utilitarianism that eschews commitment to expected utility, and avoids the problems of interpretation of scalar multiplication of policies in other arbitrage-based derivations of utilitarianism.

2. THE BOOK-MAKING ARGUMENT

A policy generates a welfare distribution over the population. Offering advanced medical treatment in an academic hospital, e.g. heart transplant, affects the health in that region relative to other regions. Consequences can be manifold and different in nature: advanced labor is attracted, earning power is affected, extra services are called for (e.g. a golf course), and house prices may rise. Offering advanced medical treatment entails a welfare distribution over the population and, hence, a policy.

In this note, the analysis is restricted to financial consequences. That is, consequences are expressed in terms of money, or can be replaced by monetary equivalents. Such replacements are, for instance, central in cost-effectiveness studies where techniques for such replacements are examined (Gold et al. 1996).

Assume a finite *population* $S = \{s_1, \dots, s_n\}$. Policies are denoted by f, g, h . The set of real numbers \mathbb{R} is the set of *consequences* and a *policy* f is a function from the population

$\{s_1, \dots, s_n\}$ to the set of consequences. The policy f describes the monetary consequence for each individual when it is implemented, with $f(s_i)$, ($i = 1, \dots, n$) the amount of money for individual s_i if policy f is implemented. Policies are often identified with n -tuples and, hence, the set of policies is identified with \mathbb{R}^n .

A social planner has a *preference relation* \succsim over the policies, with \succ (strict preference), and \sim (equivalence) as usual. We assume that \succsim is a *weak order*: \succsim is *complete* (for all policies f and g , $f \succsim g$ or $g \succsim f$, or both) and transitive.

Policies with monetary consequences can be combined, i.e., taken together at the same time in the sense of coordinate-wise addition of the policies. A social planner combines and implements several decisions taken in different domains of interest. The coordinate-wise addition assumes that there are no systematic interactions between different policies. Policies that do interact, can be combined into one policy, and policies are described through their net effect on a given situation. Such procedures for modeling decisions are common in consequentialistic approaches (Broome, 1991; Hammond, 1988; Kahneman et al., 1997).

A *Dutch book* means the following: suppose that there are two arrays of policies f^1, \dots, f^m and g^1, \dots, g^m . The superscript $j = 1, \dots, m$ represents the different domains of interest. Each policy f^j is weakly preferred to g^j by the social planner but, taking all the decisions together, i.e., combining all policies at the same time, each member of the society is worse off. The formal definition is as follows.

DEFINITION 1. A Dutch book consists of two arrays of policies f^1, \dots, f^m and g^1, \dots, g^m such that $f^j \succsim g^j$ for all policies f^j, g^j , $j = 1, \dots, m$, but $\sum_{j=1}^m f^j(s_i) < \sum_{j=1}^m g^j(s_i)$ for all individuals s_i .

Details are as follows. The social planner has to choose, for instance, between the two policies f^1 and g^1 , in the domain of interest no. 1. In practice, the social planner must be able to make her decision in domain no. 1 to a large

extent independently from the other domains. Suppose $f^1 \succcurlyeq g^1$: the social planner weakly prefers to implement policy f^1 rather than g^1 . Assume that similar preferences hold for the m different domains of interest: $f^j \succcurlyeq g^j$ for all $j = 1, \dots, m$. The sum of all the monetary consequences for individual s_i , when all the m policies are implemented, is: $f_{s_i}^1 + \dots + f_{s_i}^m$. A similar sum is considered for any member of the society s_i : $\sum_{j=1}^m f^j(s_i)$. A Dutch book means $\sum_{j=1}^m f^j(s_i) < \sum_{j=1}^m g^j(s_i)$ for all s_i .

Coherence considerations motivate the idea that if all the m policies f^j are weakly preferred to the g^j s, then one expects that when the social planner implements the policies, the population should not be worse off. This is, however, not the case when there is a Dutch book. The *book-making argument* (also called *coherence*) for social welfare is thus natural: a Dutch book should not exist.

A *constant equivalent* for a policy f is a consequence c such that (c, \dots, c) is equivalent to the policy. The concept of a constant equivalent is commonly used in the welfare literature (see Atkinson, 1970; Kolm, 1969). Next, the central result of this note is stated: the book-making theorem for welfare. The result and its proof adapt Theorem 2 of Diecidue and Wakker (2002) to welfare.

THEOREM 2. *The following two statements are equivalent for \succcurlyeq on \mathbb{R}^n .*

- (i) There exist unique nonnegative weights w_1, \dots, w_n summing to one such that policies are evaluated by $f \mapsto w_1 f(s_1) + \dots + w_n f(s_n)$.
- (ii) The binary relation \succcurlyeq is a weak order, for each policy there exists a constant equivalent, and no Dutch book can be made. □

The theorem presents a condition for the existence and the uniqueness of nonnegative weights w_j such that the social preferences over policies are represented by a welfare function that assigns a real valued index to any conceivable policy. This index is the weighted sum of the individuals'

consequences. *Harsanyi's utilitarianism* similarly states that a weighted average of individual utilities is taken as the measure of social welfare. Theorem 2 presents a foundation for utilitarianism that is alternative to Harsanyi's. Both the approaches of Harsanyi and of this paper assume linear utility, Harsanyi's with respect to probability mixtures which requires an expected utility assumption, and the present approach with respect to monetary consequences. Linear utility is reasonable for everyday decisions (Rabin, 2000) and is acceptable in utilitarianism (Neuefeind and Trockel, 1995).

The interest for Theorem 2 lies in its welfare interpretation and in its implications for utilitarianism, not in its mathematical generality. It can be demonstrated that the Pareto principle is implied by the book-making argument. The *Pareto principle* states that for all policies f and g , if the consequence under f is better than the consequence under g for every individual, then the policy f is preferred to the policy g . In symbols: for all policies f and g , if $f(s_i) > g(s_i)$, $i = 1, \dots, n$, then $f \succ g$. In the axiomatic treatment of social choice and welfare, the Pareto principle is a natural requirement, and it is not discussed further.

In the literature in economics and decision under uncertainty (Bunn, 1984; Nau and McCardle, 1990; Turunen-Red and Woodland, 1999; Varian, 1987; Yaari, 1985), the book-making argument is derived from theorems of the alternative and always invokes scalar multiplication of policies such as in $\lambda_1 f^1 + \dots + \lambda_n f^n$. In the present note, the requirements are weaker than in the mentioned contributions: only the sum, i.e., the combination of given policies is considered, which is more natural in welfare economics. In reality, if the social planner carries out a number of policies, members of society experience the combination of the effects of the policies, which is considered in Theorem 2. Individuals do not experience (combinations of) scalar multiples of the policies, which is what the existing theorems consider. In this sense, Theorem 2 is more realistic and fundamental for welfare applications.

3. EXAMPLES

We discuss two particular examples of welfare choice to clarify the book-making argument. Consequences are expressed in dollars.

EXAMPLE 3. Suppose that the social planner chooses among policies guided by equity considerations. In the domain of interest no. 1 she decides $(10, 10, 10) \succ (40, 0, 0)$, in the domain of interest no. 2 she decides $(10, 10, 10) \succ (0, 40, 0)$, and for domain of interest no. 3 she decides $(10, 10, 10) \succ (0, 0, 40)$. Taking all the decisions together: $(30, 30, 30)$ is chosen instead of $(40, 40, 40)$. A Dutch book has resulted.

EXAMPLE 4. Suppose that the social planner follows the Rawlsian maximin principle. In the domain no. 1 she decides $(10, 10, 10) \succ (0, 30, 30)$, in the domain of interest no. 2 she decides $(10, 10, 10) \succ (30, 0, 30)$, and in the domain of interest no. 3 she decides $(10, 10, 10) \succ (30, 30, 0)$. Implementing all the policies yields: $(30, 30, 30)$ instead of $(60, 60, 60)$. A Dutch book has resulted.

It is obvious that, since a Dutch book has resulted in both examples, equity and Rawlsian maximin (Rawls, 1971) when applied in myopic manners in different domains of interest, are not compatible with Theorem 2, and myopic applications can harm the society.

A prescriptive requirement is that, once decisions are aggregated (for example at the national level), the society as a whole should not be harmed: thus no Dutch book should arise. A more general policy implication for equity considerations: policies should be considered and evaluated at an aggregated level. Non-integrated equity considerations can harm society.

4. CONCLUSION

Utilitarianism is the most popular model in applied policy analysis for welfare. The classic Harsanyi's approach to

utilitarianism relies on expected utility and decision under risk: This is somewhat unnatural in welfare economics. Further, expected utility is violated empirically and its normative status has also been challenged (Allais, 1953; Camerer, 1995; Machina, 1989; Rabin, 2000).

This note has presented a new foundation for utilitarianism alternative to Harsanyi's and based on de Finetti's book making-argument (1931). The argument entails that a Dutch book should not exist. De Finetti's condition is nowadays a standard argument in the theory of subjective probabilities, with its own esthetic value based on its pragmatism. The appeal of the book-making argument for welfare is given by the simple economic intuition underlying it. To aggregate preferences in a linear manner, the following is requested: a weak order, the constant equivalent, and no Dutch book allowed. This last assumption has a natural interpretation and compelling implications for policy making and welfare economics.

This note gives a new interpretation of the original book-making argument as a coherence principle for welfare decisions. By avoiding the strong assumptions of the classic Harsanyi's result it provides a cleaner, new foundation of utilitarianism. It is remarkable that Harsanyi's famous defense of utilitarianism and de Finetti's famous book-making argument are such close relatives, both being based on linear utility.

ACKNOWLEDGEMENTS

I wish to thank Han Bleichrodt, Peter P. Wakker, and John Weymark for helpful suggestions.

REFERENCES

- Allais, M. (1953), The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American school, in Allais, M. and Hagen, O. (1979, eds.), *Expected Utility Hypotheses and the Allais Paradox*, Dordrecht: Reidel, pp. 27–145.

- Atkinson, A.B. (1970), On the measurement of inequality, *Journal of Economic Theory* 2, 244–263.
- Broome, J.R. (1991), *Weighing Goods*, Oxford, UK: Basil Blackwell.
- Bunn, D.W. (1984), *Applied Decision Analysis*, New York: McGraw-Hill book Company.
- Camerer, C. (1995), Individual decision making, in Kagel, J.H., and Roth, A.E. (eds.), *Handbook of Experimental Economics*, Princeton, NJ: Princeton University Press, pp. 587–703.
- de Finetti, B. (1931), Sul significato soggettivo della probabilità, *Fundamenta Mathematicae* 17, 298–329. Translated into English as ‘On the subjective meaning of probability,’ in Monari, P. and Cocchi, D. (eds., 1993), *Probabilità e Induzione*, Bologna: Clueb, pp. 291–321.
- Diecidue, E. and Wakker, P.P. (2002), Dutch books: avoiding strategic and dynamic complications, and a comonotonic extension, *Mathematical Social Sciences*, 43, 135–149.
- Ebert, U. (1988), Measurement of Inequality: an attempt at unification and generalization, *Social Choice and Welfare* 5, 147–169.
- Gold, M.R., Siegel, J.E., Russell, L.B. and Weinstein, M.C. (1996), *Cost-Effectiveness in Health and Medicine*, New York: Oxford University Press.
- Hammond, P.J. (1988), Consequentialist foundations for expected utility, *Theory and Decision* 25, 25–78.
- Harsanyi, J.C. (1955), Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of utility, *Journal of Political Economy* 63, 309–321.
- Kahneman, D. and Tversky, A. (1979), Prospect theory: an analysis of decision under risk, *Econometrica* 47, 263–291.
- Kahneman, D., Wakker, P.P. and Sarin, R.K. (1997), Back to Bentham explorations of experienced utility, *Quarterly Journal of Economics* 112, 375–405.
- Kolm, S.C. (1969), The optimal production of social justice, in Margolis, J. and Guitton, H. (eds.), *Public Economics*, London: MacMillan, pp. 145–200.
- Machina, M.J. (1989), Dynamic consistency and non-expected utility models of choice under uncertainty, *Journal of Economic Literature* 27, 1622–1688.
- Nau, R.F. and McCardle K.F. (1990), Coherent behavior in noncooperative games, *Journal of Economic Theory* 50, 424–444.
- Neufeld, W. and Trockel, W. (1995), Continuous linear representability of binary relations, *Economic Theory* 6, 351–356.
- Rabin, M. (2000), Risk aversion and expected-utility theory: a calibration theorem, *Econometrica* 68, 1281–1292.
- Rawls, J. (1971), *A Theory of Justice*, Cambridge, MA: Harvard University Press.
- Savage, L.J. (1954), *The Foundations of Statistics*, New York: Wiley.

- Schmidt, U. (2004), Alternatives to expected utility: some formal theories, in Barberà, S., Hammond, P.J. and Seidl, C. (eds.), *Handbook of Utility Theory*, Vol. 2, Dordrecht: Kluwer Academic Publishers, pp. 757–838.
- Starmer, C. (2000), Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk, *Journal of Economic Literature* 38, 332–382.
- Turunen-Red, A.H. and Woodland, A.D. (1999), On economic applications of the Kuhn-Fourier Theorem, in Wooders, M.H. (ed.), *Topics in Mathematical Economics and Game Theory: Essays in Honor of Robert J. Aumann*, Toronto: University of Toronto.
- Varian, H.R. (1987), The arbitrage principle in financial economics, *Journal of Economic Perspectives* 1(2), 55–72.
- von Neumann, J. and Morgenstern, O. (1947), *Theory of Games and Economic Behavior*, Princeton NJ: Princeton University Press.
- Weymark, J.A. (1981), Generalized gini inequality indices, *Mathematical Social Sciences* 1, 409–430.
- Yaari, M.E. (1985), On the role of “Dutch books” in the theory of choice under risk. 1985 Nancy L. Schwartz Memorial Lecture, Evanston, IL: J. L. Kellogg Graduate School of Management, Northwestern University.
- Yaari, M.E. (1988), A controversial proposal concerning inequality measurement, *Journal of Economic Theory* 44, 381–397.

Address for correspondence: Enrico Diecidue, INSEAD, Boulevard de Constance, 77305, Fontainebleau Cedex, France.
E-mail: enrico.diecidue@insead.edu