Deriving Harsanyi's Utilitarianism from de Finetti's Book-Making Argument

Enrico Diecidue*

INSEAD, Fontainebleau and Singapore

June 9, 2003

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. In this note, a similar argument is applied to a welfare context and its implications are studied. The new argument provides a foundation for utilitarianism that is alternative to Harsanyi's, and that generalizes foundations based on the theorem of the alternative.

Keywords: Book-making argument, utilitarianism, policy independence

JEL Classification Number: D60, D71, D81

*Address: INSEAD, Boulevard de Constance, 77305 Fontainebleau Cedex, France.

enrico.diecidue@insead.edu

1 Introduction

This note presents a coherence condition for welfare evaluations. It is based on the book-making argument introduced by de Finetti [6], 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 [16] 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 [8] by avoiding reliance on expected utility. It generalizes some other existing foundations of utilitarianism based on the theorem of the alternative by using only natural combinations of policy decisions, and not scalar multiplications.

2 The book-making argument

Assume a finite *population* $S = \{s_1, ..., s_n\}$. Informally, a policy generates a welfare distribution over the population. Policies are denoted by f, g, h. They can have different kinds of consequences for the population: financial, sociological, environmental, health- or security-oriented, etc. Building a new plant for the treatment of medical waste material, for example, entails a welfare distribution over the population and, hence, a policy. Its consequences can be manifold and different in nature:

the emissions of the new plant can affect the environment and thus the well-being of the population, the health of people living in the neighborhood of the plant can deteriorate, the price of rent and housing can decrease, and so on.

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. Formally, the set of real numbers $I\!R$ is the set of *consequences* and a *policy* f is a function from the population $\{s_1, ..., s_n\}$ to the set of consequences. The policy f describes the monetary consequence for each individual when it is implemented, with $f(s_i)$, $(i = 1, ..., n)$ 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 $I\!R^n$.

A social planner (or a government) has a *preference relation* $\succcurlyeq$ over the policies, with $\succ$ (strict preference) and $\sim$ (equivalence) as usual. We assume that $\succcurlyeq$ is a *weak order*: $\succcurlyeq$ is *complete* (for all policies f and g, $f \succcurlyeq g$ or $g \succcurlyeq f$, 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 several decisions taken in different areas of interest. A government combines and implements the decisions taken by the responsible ministers. The coordinate-wise addition assumes that there are no systematic interactions between different policies. Policies that do interact, should be combined into one policy, and policies are described through their net effect on a given situation.

The novelty of this note is to present the book-making argument for welfare.

The argument permits to naturally justify Harsanyi's utilitarianism in an alternative manner. The analysis is formulated for the case of a government as the decision maker. Similar considerations apply to other cases.

A *Dutch book* means the following: suppose that there are two arrays of policies $f^1, ..., f^m$ and $g^1, ..., g^m$. Each policy f^j is weakly preferred to g^j by the responsible minister 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* $f^1, ..., f^m$ *and* $g^1, ..., g^m$ *such that* $f^j \succcurlyeq g^j$ *for all policies* f^j, g^j, $j = 1, ...m$, *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. Minister #1 has to choose, for instance, between the two policies f^1 and g^1. The minister chooses independently from the other minister's decisions. Suppose $f^1 \succcurlyeq g^1$: minister #1 weakly prefers to implement policy f^1 rather than g^1. Assume that similar preferences hold for the m ministers: $f^j \succcurlyeq g^j$ for all $j = 1, ..., m$. The sum of all the monetary consequences for individual s_i, when all the m policies are implemented, is: $f^1_{s_i} + ... + f^m_{s_i}$. 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^js, then one expects that when the government implements the policies, the population should not be worse off. This is, however, not the case when there is a Dutch book. If a Dutch book as in Definition 1 exists, then every member of

the society is strictly worse off in monetary terms, a situation that should obviously be avoided. 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 $(c, ..., c)$ is equivalent to the policy. The concept of a constant equivalent is commonly used in the welfare literature, see Kolm [10] and Atkinson [2]. Next, the central result of this note is stated: the book-making theorem for welfare. It adapts Theorem 2 of Diecidue & Wakker [7] to welfare.

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

(i) There exist unique nonnegative weights $p_1, ..., p_n$ summing to one such that policies are evaluated by $f \mapsto p_1 f(s_1) + ... + p_n f(s_n)$.

(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 p_j such that the social preferences over policies are represented by 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. Expected utility has been criticized by many authors (Allais [1], Machina [11], Camerer [5], Rabin [14]). Linear utility is reasonable for everyday decisions (Rabin [14]) and is acceptable in utilitarianism (Neuefeind & Trockel [13]).

The interest for Theorem 2 lies in its welfare interpretation and in its implications for utilitarianism, not in its mathematical generality. The proof is based on Theorem 2 in Diecidue & Wakker [7] with a few modifications to adapt it to a welfare framework. It can be demonstrated, and is an immediate corollary of the theorem, that the Pareto principle is implied by the book-making argument. The *Pareto principle* states: for all policies f and g, if $f(s_i) > g(s_i)$, $i = 1, ..., n$, 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.

Many contributions to the book-making argument are available in decision under uncertainty, social choice, game theory, finance, etc. (Bunn [4], Nau & McCardle [12], Turunen-Red & Woodland [17], Varian [18], Yaari [19]). In all these works the book-making argument is derived from theorems of the alternative and always invokes scalar multiplication of policies such as in $\lambda_1 f^1 + ... + \lambda_n f^n$. 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 reality, if the government carries out a number of policies, every member of the society experiences the combination of the effects of the policies, which is what 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 three ministers choose among policies guided by equity considerations. Minister 1 decides* $(10, 10, 10) \succcurlyeq (40, 0, 0)$*, minister 2 decides* $(10, 10, 10) \succcurlyeq (0, 40, 0)$*, and minister 3 decides* $(10, 10, 10) \succcurlyeq (0, 0, 40)$*. Taking all the decisions together:* $(30, 30, 30)$ *is chosen instead of* $(40, 40, 40)$*. A Dutch book has resulted.*

Example 4 *Suppose that three ministers follow the Rawlsian maximin principle. Minister 1 decides* $(10, 10, 10) \succcurlyeq (0, 30, 30)$*, minister 2 decides* $(10, 10, 10) \succcurlyeq (30, 0, 30)$*, and minister 3 decides* $(10, 10, 10) \succcurlyeq (30, 30, 0)$*. 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 elementary examples, equity and Rawlsian maximin (Rawls [15]) when applied in myopic manners by subparts of the government, 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. 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

This note presents a coherence argument for welfare entailing that a Dutch book should not exist. De Finetti's [6] 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 in welfare economics. This note gives a new interpretation of the original book-making argument as a coherence principle for welfare decisions. This re-interpretation avoids the strong assumptions of the classic Harsanyi's result and in this sense 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.

References

[1] 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 M. Allais and O. Hagen (1979, Eds.), *Expected Utility Hypotheses and the Allais Paradox* (Reidel, Dordrecht), 27-145.

[2] Atkinson, A.B. (1970) "On the Measurement of Inequality" *Journal of Economic Theory* 2, 244-263.

[3] Buehler, R. J. (1976) "Coherent Preferences," *Annals of Statistics* 4, 1051-1064.

[4] Bunn, D. W. (1984) *Applied Decision Analysis*. McGraw-Hill book Company, New York.

[5] Camerer, C. (1995) "Individual Decision Making." In John H. Kagel & Alvin E. Roth (Eds.), *Handbook of Experimental Economics*, 587-703, Princeton University Press, Princeton, NJ.

[6] 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 Paola Monari and Daniela Cocchi (Eds., 1993) "*Probabilità e Induzione*," Clueb, Bologna, 291-321.

[7] Diecidue, E., and P. P. Wakker (2002) "Dutch Books: Avoiding Strategic and Dynamic Complications, and a Comonotonic Extension" *Mathematical Social Sciences*, 43, 135-149.

[8] Harsanyi, J. C. (1955) "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility" *Journal of Political Economy* 63: 309-321.

[9] Kelsey, D. (1995) "Dutch Books Arguments and Learning in a Nonexpected Utility Framework" *International Economic Review*, 36: 187-286.

[10] Kolm, S. C. (1969) "The Optimal Production of Social Justice" in J. Margolis and H. Guitton, eds., *Public Economics* (MacMillan, London), 145-200.

[11] Machina, M. J. (1989) "Dynamic Consistency and Non-Expected Utility Models of Choice under Uncertainty" *Journal of Economic Literature* 27, 1622-1688.

[12] Nau, R. F., and K. F. McCardle (1990) "Coherent Behavior in Noncooperative Games" *Journal of Economic Theory* 50, 424-444.

[13] Neuefeind, W., and W. Trockel (1995) "Continuous Linear Representability of Binary Relations" *Economic Theory* 6: 351-356.

[14] Rabin, M. (2000) "Risk Aversion and Expected-utility Theory: A Calibration Theorem," *Econometrica* 68, 1281-1292.

[15] Rawls, J. (1971) *A Theory of Justice*, Harvard University Press, Cambridge, MA.

[16] Savage, L. J. (1954) *The Foundations of Statistics*. Wiley, New York. (Second edition 1972, Dover, New York.)

[17] Turunen-Red, A. H., and A. D. Woodland (1999) "On Economic Applications of the Kuhn-Fourier Theorem" In Myrna H. Wooders (Ed.), *Topics in Mathematical*

Economics and Game Theory: Essays in Honor of Robert J. Aumann. University of Toronto.

[18] Varian, H. R. (1987) "The Arbitrage Principle in Financial Economics" *Journal of Economic Perspectives* 1 no. 2: 55-72.

[19] Yaari, M. E. (1985) "On the Role of "Dutch Books" in the Theory of Choice under Risk" 1985 Nancy L. Schwartz memorial lecture, J. L. Kellogg Graduate School of Management, Northwestern University