

Interpersonal Comparisons and Utilitarian Social Choice

by John Lawrence

Abstract

Much has been written about the irrelevance of utilitarian approaches to social choice because there is supposedly no way to compare interpersonal measures of utility. In Kenneth Arrow's classic book, "Social Choice and Individual Values,"¹ he rejected the utilitarian approach because of this arbitrariness. Rather than individual preference ratings consisting of points on a real number utility scale, his starting point was individual preference rankings, $aPbPc$, a is preferred to b is preferred to c etc. This approach has effectively eliminated considerations of utility in social choice theory. Arrow then went on to show that there was no possible way that preference rankings could be combined to produce a social ranking such that certain reasonable criteria were met. This conclusion proved that social choice was impossible. Gibbard² and Satterthwaite³, concurring with Arrow, showed that any social choice system that was strategy proof was impossible as well. However, as we prove in this paper, there is at least one example of a social choice system in which the outcome for society can be based on utility theory in such a way that Arrow's and Gibbard-Satterthwaite's Impossibility Theorems don't apply. This method, which we call Optimal Threshold Social Choice (OTSC), is a hybrid of utilitarian and approval voting that satisfies Arrow's criteria, is strategy proof and makes interpersonal comparisons moot. It applies to multi-winner elections and also to the selection of consumer bundles in the theory of consumer choice. We show in this paper that the outcomes of OTSC, which are based on utilitarian input information, don't depend on the scale on which utility is measured. Furthermore, Arrow's method of using preference rankings can be, without loss of generality, converted into utilitarian preference ratings as inputs to the OTSC and processed in such a way that bonafide social choice outcomes including rankings are produced. This effectively negates both impossibility theorems.

Introduction

In "Social Choice and Individual Values", Arrow states⁴: "If we admit meaning to interpersonal comparisons of utility, then presumably we could order social states according to the sum of utilities of individuals under each, and this is the solution of Jeremy Bentham, accepted by Edgeworth and Marshall." It has long been thought⁵ that utilitarian voting meets Arrow's five criteria for a rational voting system although it is not strategy proof. It is arguable whether or not utilitarian voting violates Arrow's Condition of Independence of Irrelevant Alternatives. We assume it does not. However, it fails Gibbard-Satterthwaite's Impossibility criterion because voters may vote insincerely in order to gain an advantage. Arrow's solution to the social choice conundrum is to disregard considerations of utility altogether. He states⁶: "The viewpoint will be taken here that interpersonal comparison of utilities has no meaning and, in fact, that there is no meaning relevant to welfare comparisons in the measurability of individual utility." So Arrow sets out to find a social choice mechanism which is not utility based because comparison of individual utilities is meaningless. Still there must be a relationship between any social choice solution and utility. Choosers want the social choice to maximize, or at least have a relationship to, their utilities in some sense; otherwise, the solution is disconnected from real life. We take the approach of devising a method that is utility based but whose final results are not based on the sum of utilities.

The Purpose of Social Choice

We want the candidate set whether it's a President or a Parliament, a Congress or a Senate to represent

us, be favorable to us. Likewise, we want our selection of consumer items to give us satisfaction. So individual choice is utility based whether or not that is acknowledged by theoreticians. Consequently, social choice should also reflect individual utilities. What exactly does that mean? For me it means that it reflects my satisfaction with a set of candidates or a set of consumer bundles. For instance, I can rate on some scale how much I like each candidate that is running for some position. Let's arbitrarily pick a scale from one to ten. Then if I rate someone a 10 that means I have maximum utility for that candidate. Conversely, if I rate someone a 1, that means I have minimum utility for that candidate. Utility can also mean "happiness with" or "approval of." People are asked to quantify their feelings about some issue all the time so this should not come as a shock to most people. For instance, a doctor will ask you to quantify your pain on a scale from one to ten with ten being the most severe pain. So to quantify your satisfaction for some individual or some basket of consumer items should not be a problem in theory. Instead of a scale from one to ten we could use a scale from minus one to plus one or a scale from 0 to 100 or any other scale. Arrow's point is that the scale itself is arbitrary, but, as shown in this paper, any affine linear transformation of a given utility scale will yield the same results for the social choice system under consideration.

In general a social choice system can be thought of as having an input from the voters or choosers which is unconstrained as to whether it is preference rankings, ratings or some other form of input, then a processing unit which is a much more general process than a social ranking, rating or "sum of utilities." Finally, the output gives the results of the process which may or may not be unique to each individual as in the following diagram.



Utilitarian Voting

Arrow concedes that⁷ "... for each individual, the utility indicator can be defined uniquely among the previously defined class, which is unique up to a linear transformation, by assigning the utility 1 to the best conceivable alternative and 0 to the worst conceivable alternative. This assignment of values is designed to make individual utilities interpersonally comparable." Arrow then shows an example in which the individual utilities change depending on the the composition of the alternative set thereby violating his Condition 3, Independence of Irrelevant Alternatives. However, in utilitarian voting and range voting⁸ in particular, individual utilities are assumed not to change as a function of the composition of the alternative set. They remain fixed and, therefore do satisfy Arrow's Condition 3. Therefore, utilitarian voting satisfies Arrow's conditions. However, utilitarian voting does violate the Gibbard-Satterthwaite Impossibility theorem because it is not strategy proof. A well known strategy in an election, the outcome of which returns one candidate, is to give the favorite candidate a 1 and all the rest of the candidates 0s even though some other candidates may have finite utilities. Thus, insincere voting is possible in range or utilitarian voting.

Strategyproofness

We have shown elsewhere⁹ that strategyproofness for individual choosers can be gained by making strategy part of the social choice processing unit itself. The individual chooser is disincentivized from using strategy since the system that processes the individual choices does it in such a way as to maximize the power of each individual choice, thus giving each individual the optimum strategy. For an individual to use strategy in the submission of their choices would only diminish the outcome produced for them individually by the system. Therefore, the individual chooser is incentivized to choose sincerely. The optimum strategy is to set a threshold in each individual's utilitarian style input which gives every candidate above threshold the maximum “vote” and every candidate below threshold the minimum “vote” in such a way as to maximize the expected value of utility of the social choice for each individual. This effectively turns the utilitarian style input into approval style “votes,” but the connection with the utilitarian basis of the system is maintained. The following diagram indicates how the choosing system, in this case, OTSC, works.



The details for the Optimal Threshold Filter are found in my paper¹⁰, “Optimal Threshold for Selection of Candidates in Multi-Winner Elections.” The optimum strategy for each voter is to vote for or choose those candidates which maximize the average of the voter's above threshold utilities and reject all the rest. This determines where the threshold is to be placed for each individual voter. By taking this procedure out of each individual's hands and placing it in the processing system itself, the individual is disincentivized from voting insincerely.

Interpersonal Comparisons Are Moot for OTSC

Arrow¹¹ dwells on the fact that individual utility scales are not compatible. He compares them with the measurement of temperature which is based on arbitrary units and the arbitrary terminal points of freezing and boiling for the Celsius scale and completely different end points for the Fahrenheit scale. “Even if, for some reason, we should admit the measurability of utility for an individual, there still remains the question of aggregating the individual utilities. At best, it is contended that, for an individual, his utility function is uniquely determined up to a linear transformation; we must still choose one out of the infinite family of indicators to represent the individual, and the values of the aggregate (say a sum) are dependent on how the choice is made for each individual. In general, there seems to be no method intrinsic to utility measurement which will make the choice compatible.”

Murakami¹² also considers this problem of arbitrariness:

“However, the origin and unit of measurement of preference intensity is, as we have noted, arbitrary. As we choose the Fahrenheit system or the Centigrade system in measuring temperature, so we have to determine the origin and unit of measurement here. In other words we have to select two 'base' alternatives, such as freezing point and boiling point in the Centigrade system, for which the magnitudes of utility are fixed.” Murakami favors a cardinal utility index which is based on a utilitarian scale with two fixed points analogous to the freezing and boiling points on the Centigrade scale.

Harsanyi¹³ argues that the issue of interpersonal comparisons can be resolved by considering some kind of average of different individuals' utility scales. "In any case it should now be sufficiently clear that interpersonal comparisons of utility are not value judgments based on some ethical or political postulates, but rather are factual propositions based on certain principles of inductive logic." He concedes that each individual has a unique utility scale. The major point of the present paper is that, regardless of each individual's utility scale, regardless of the point of origin or points of reference, the Optimal Threshold filter or processing unit will yield the same results. This is proven in Appendix 1. OTSC maximizes the power of each individual vote or choice regardless of the individual utility scale that any individual uses. At the same time it disincentivizes any insincere voting or choosing.

Bonner¹⁴ has discussed cardinal utility as follows: "Cardinal measurement is of little use in adding up social welfare if interpersonal comparisons cannot be made. ... The scale and origin of every personal index might be different, and – what is more important – any attempt to convert them to a common basis would be open to criticism." We have shown that even though the scale and origin of every personal index may be different, the OTSC method can process them in such a way that each individual's input will yield maximal results for them. Regardless of any affine linear transformation of each utility scale, the results for OTSC will be the same so that the individual chooser is free to choose any scale he or she wants. Since this is true, any convenient scale such as the real line between -1 and +1 can be chosen without any loss of generality or arbitrariness. Or a range of choices between the integers 0 and 99 can be chosen as in range voting⁴.

Let's unpack Bonner's statement. First, let's admit the measurability of utility for each individual. Let's say that, in general, utility can be measured as points on the real line where $-\infty < x < +\infty$ and x is a point of the real line. It's up to the individual where to place the points, including the end points, corresponding to the utilities of each candidate in the candidate set consisting of m candidates, $\{c_1, c_2, \dots, c_m\}$ where the candidates could be either individuals running for election or baskets of consumer items. It is proven in the appendix that, for the OTSC system in particular, the results will be the same no matter which utility scale each individual chooses. There is no need to "choose one out of the infinite family of indicators to represent the individual." Consequently, Arrow's statement that "the values of the aggregate are dependent on how the choice is made for each individual" is not necessarily true.

The OTSC procedure converts an individually specified set of utilities regardless of scale to a set of approval style decisions (ones and zeroes, or alternatively, ones and minus ones). The ones represent the "votes" for candidates in the candidate set. This conversion is done in such a way as to maximize the power of each individual vote. Therefore, the choice made for each individual is "compatible" since it's made using the same rationale. It has been shown that OTSC is strategy proof⁵. No matter which scale an individual chooses, he or she has no incentive to misrepresent their true utilities.

Let us consider how the OTSC method compares with the sum of utilities method. If all individual utilities were normalized such that the candidate with the greatest utility was placed at "1" on the utility scale, the candidate with the least utility was placed at "0", for example, and the utilities for each candidate were added up, what would the comparison show? It can be conjectured that the sum of utilities method, if each chooser chose sincerely, would result in greater overall utility for society. This assumes that each chooser is given the same utility scale thereby democratizing the choosing power of each. However, there might be a price to be paid for the strategy proofness of OTSC, and that price might result in a diminution of social utility compared to the sum-of-utilities method. Alternatively, the Bayesian regret¹⁵ of the OTSC system can be compared to that of other systems including an ideal system.

Amartya Sen stated in his Nobel lecture¹⁶ "... economists came to be persuaded by arguments presented by Lionel Robbins and others (deeply influenced by "logical positivist" philosophy) that interpersonal comparisons of utility had no scientific basis. 'Every mind is inscrutable to every other mind and no common denominator of feelings is possible.' Thus, the epistemic foundations of utilitarian welfare economics were seen as incurably defective." OTSC has shown that there is a sound epistemic basis for a utility based social choice system, and, therefore, the OTSC system is in fact logical positivist.

Preference Rankings Can Be Converted to Ratings and Processed by OTSC

Arrow's preference rankings can be converted to utility scales for each individual which are then passed through the same OTSC procedure. Since the only information for rankings is of the form $aPbPcPd...$ which is interpreted as a is preferred to b , b is preferred to c etc., we can choose any utility scale as long as the preference ratings are equally spaced since that is the only information we have. We know that the choice of which scale to use is irrelevant. Let's say we choose the real line between 0 and 100. We let the top ranked candidate be placed at 100 and the lowest ranked candidate be placed at 0. The other candidates are equally spaced on the scale. Then, since an optimal threshold exists, the OTSC information processing system outputs approval style positive choices for those candidates represented by utilities above threshold and negative choices for those candidates represented by utilities below threshold. As we have shown, any affine linear transformation will not change the results of the OTSC processing system. Obviously, the outputs are in the form of integers. If output integers are $x > y > z$, for example, then $xPyPz$ is the social choice. This conversion from rankings to ratings can be made part of the processing system itself, and the social choice inputs and outputs will be exactly in the form Arrow assumed. Therefore, Arrow's Impossibility Theorem is disproven.

Conclusions

We have shown that the issue of interpersonal comparisons in social choice theory is moot at least for a particular information processing system that we call Optimal Threshold Social Choice (OTSC)¹⁰. Any affine linear transformation of an individual's utility scale will produce identical results. It is known that utilitarian or range voting satisfies Arrow's criteria, and, therefore, Arrow's Impossibility Theorem is reversed if preference ratings instead of preference rankings are admitted as input. However, utilitarian or range voting is not strategyproof. It has been shown previously that OTSC is strategyproof⁵, and, therefore, the Gibbard-Satterthwaite Theorem is also reversed. We have demonstrated a utilitarian based social choice system for which not only is the interpersonal comparison issue irrelevant but also satisfies Arrow's criteria, and, furthermore, is strategy proof and can even accept preference rankings as input and produce social rankings as output exactly as in Arrow's proposed model.

Appendix

To prove: Given any arbitrary individual utility scale consisting of preference ratings as inputs, the social choice results, when processed by the OTSC, will be the same as they would be for any affine linear transformation of that scale.

For some particular voter let the expected value of above threshold average utility be

$$(p_2u_2 + \dots + p_nu_n)/n$$

where $p_j u_j < p_{j+1} u_{j+1}$ for $2 \leq j \leq n-1$ and the optimal threshold is just under $p_2 u_2$. We perform a linear affine transformation of the form $f(x) = ax + b$ and assume that the optimal threshold will move down from just under $p_2 u_2$ to just under $p_1 u_1$ so that the above threshold average utility is now $(p_1 u_1 + \dots + p_n u_n)/n$

We assume:

$$\frac{\sum_1^n (a u_j + b)}{n} > \frac{\sum_2^n (a u_j + b)}{n-1}$$

We know:

$$\frac{\sum_1^n (a u_j + b)}{n} = \frac{\sum_1^n (a u_j) + n b}{n} = \frac{a \sum_1^n (u_j)}{n} + b$$

$$\frac{\sum_2^n (a u_j + b)}{n-1} = \frac{\sum_2^n (a u_j) + (n-1)b}{n-1} = \frac{a \sum_2^n (u_j)}{n-1} + b$$

So is

$$\frac{a \sum_1^n (u_j)}{n} + b > \frac{a \sum_2^n (u_j)}{n-1} + b$$

Subtracting b from both sides and dividing by a we have

$$\frac{\sum_1^n (u_j)}{n} > \frac{\sum_2^n (u_j)}{n-1}$$

However, we know that

$$\frac{\sum_1^n (u_j)}{n} < \frac{\sum_2^n (u_j)}{n-1}$$

because by definition the optimal threshold is placed just under the utility such that the average utility above threshold is a maximum.

QED

Similarly, if the average above threshold utility is $(p_1 u_1 + \dots + p_n u_n)/n$, we show that applying an affine linear transformation and assuming that the optimal threshold moves up to just under $p_2 u_2$ is false.

Assume that

$$\frac{\sum_2^n (au_j + b)}{n-1} > \frac{\sum_1^n (au_j + b)}{n}$$

Then

$$\frac{\sum_2^n (au_j + b)}{n-1} = \frac{\sum_2^n (au_j) + (n-1)b}{n-1} = \frac{a \sum_2^n (u_j)}{n-1} + b$$

and

$$\frac{\sum_1^n (au_j + b)}{n} = \frac{\sum_1^n (au_j) + nb}{n} = \frac{a \sum_1^n (u_j)}{n} + b$$

Therefore,

$$\frac{a \sum_2^n (u_j)}{n-1} + b > \frac{a \sum_1^n (u_j)}{n} + b$$

and

$$\frac{a \sum_2^n (u_j)}{n-1} > \frac{a \sum_1^n (u_j)}{n}$$

$$\frac{\sum_2^n (u_j)}{n-1} > \frac{\sum_1^n (u_j)}{n}$$

But we know that,

$$\frac{\sum_1^n (u_j)}{n} > \frac{\sum_2^n (u_j)}{n-1}$$

by definition of the optimal threshold and the assumption is false. QED.

Therefore, an affine linear transformation does not change the placement of the optimal threshold

References

1. Arrow, Kenneth J. (1951), *Social Choice and Individual Values*, Yale University Press.
2. Gibbard, A. (1973), "Manipulation of voting schemes: a general result," *Econometrica*, Vol. 41, No. 4, pp. 587–601.
3. Satterthwaite, M. A. (1975), "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory*, 10, pp. 187–217.
4. Arrow, *ibid.*, p. 4.
5. <https://rangevoting.org/ArrowThm.html>
6. Arrow, *ibid.*, p.9.
7. Arrow, *ibid.*, p.32.
8. <https://rangevoting.org/ArrowThm.html>
9. Lawrence, John (2019), "Arrow's and Gibbard-Satterthwaite's Impossibility Theorems Revisited," SageSubmissions. Preprint.
10. Lawrence, John C. (2019), "Optimal Threshold for Selection of Candidates in Multi-Winner Elections," SageSubmissions. Preprint. Arrow, pp10-11
11. Arrow, *ibid.* p.10.
12. Murakami, Y. (1968), *Logic and Social Choice*, Routledge & Kegan Paul Ltd., p. 114.
13. Harsanyi, J. C. (1955), "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility," *Journal of Political Economy*, 63, pp. 309-21.
14. Bonner, John (1986), "Introduction to the Theory of Social Choice," The Johns Hopkins University Press, p. 166.
15. <https://www.rangevoting.org/BayRegDum.html>
16. Sen, Amartya (1998), "The Possibility of Social Choice," Nobel Lecture, Trinity College, Cambridge, Great Britain.