

Aggregating Rights (A Social Welfare Function Based on Human Rights)

A social welfare function (henceforth SWF) ranks conceivable or possible worlds - also called "social states" - according to the aggregate levels of welfare they produce. Each social state is given a numerical value by the SWF allowing it to be ranked. This should help policy makers and individuals to achieve the social state with the highest - or maximum - aggregate welfare. Subject of course to whatever constraints there are: technological or resource constraints, human capital or behavioral constraints etc. A "higher" rather than the "highest" welfare may be the only realistic option.

A SWF therefore describes society as a whole, but the whole is an aggregate of individual welfare scores, also called individual "utilities", "interests" or whatever. Individuals should first order possible social states from their own point of view and their own sense of utility or preference, and then these individual orderings of social states are aggregated in some way by a SWF so as to achieve maximum aggregate welfare.

A SWF typically looks like this:

$$W(x) = F(U_1(x), U_2(x), \dots, U_n(x))$$

where:

$W(x)$ is a real number social utility value for a possible world x

$U_1(x)$ is a cardinal or ordinal utility value that is interpersonally comparable and that is yielded by some policy or procedure for individual 1 in world x

F is a function that yields a real number; F can be different things: the sum of all U 's, the average of all U 's, the product (multiplication) of all U 's etc.¹

n is the total number of individuals.

SWFs are hotly contested. What should count as individual "welfare" or "utility"? And how do we measure, rank or score it? And if we have an uncontested definition of individual welfare that can be measured and ranked, then how do we compare and aggregate it across individuals? In other words: what is U and what is F ?

Let's start with the first problem. Income or wealth (counted cardinally in \$) is often cited as a good proxy for welfare. When people have more money, they can pay for more goods, services or activities which give them some amount of utility in some form and which therefore enhance their welfare. Income or wealth is also measurable and interpersonally comparable. Aggregating it into a SWF should be relatively easy. Of course, F in this case should probably not be the simple sum of all income levels. Some special focus on the income of the least well off and redistribution of income from those better off would, given the Law of Diminishing Marginal Utility, increase total welfare utility (W) in a society: one extra unit of utility (income in this case) for a starving person is of greater value than an extra unit of utility for a millionaire.²

¹ A sum could be a Benthamite or classical utilitarian SWF; the product a Bernoulli-Nash SWF.

² A [Rawlsian SWF](#) would then perhaps be an option. A [Benthamite or classical utilitarian SWF](#), even though it would allow redistribution if justified by diminishing marginal utility, does not give special attention to the income of the least well off. A Rawlsian SWF allows redistribution even without diminishing marginal utility.

However, income and wealth aren't really very good proxies for welfare. Welfare is clearly more than money, even though money can buy welfare. In addition, different people need different levels of income or wealth because they have different consumption requirements (think of a blind person or a paraplegic). So income looks like it's an interpersonally comparable form of welfare when in fact it's not: equal income may yield vastly unequal levels of welfare when welfare is understood as something more than mere income. Well-being may then have to take the place of income (well-being could account for the lower welfare of the blind and the paraplegic given equal income levels). But well-being is notoriously difficult to define, let alone measure.³ So we're back where we came from.

Perhaps a more subjective approach to welfare could do the trick. Welfare is then not what is objectively good for individuals - such as income or well-being - but rather what individuals themselves subjectively prefer. This means moving away from what Kenneth Arrow has called the Platonism or paternalism that is often part of welfare economics.⁴ The social good is not derived from philosophical reflection about the nature of welfare, but is instead the composite of individual preferences, whatever they are. Hence, the welfare that is to be aggregated into a SWF is abstract preference satisfaction.

The use of abstract preference satisfaction as a measure of individual welfare allows us to sidestep the problem of interpersonal comparability of more specific preferences or of other possible definitions of welfare such as income. There is no way to compare satisfaction of a preference for alcohol with satisfaction of a preference for a long life. Of course, even if we try to compare abstract preference satisfaction across individual, we'll still have problems because "[t]here is no means of testing the magnitude of A's satisfaction as compared with B's".⁵ Hence the need to use ordinal rather than cardinal rankings for individual preference rankings. One takes individuals' ordinal preference rankings and aggregates those, irrespective of the nature or intensity of the preferences. However, we then may have to pay the price of running afoul of [Arrow's Impossibility Theorem](#) which suggest that individual ordinal rankings can't produce a SWF.

Moreover, the interpersonal comparability of abstract preference satisfaction is only apparent. Let's say you have a world with three individual, each with a set of preferences that they rank. Two of them have a high preference for the consumption of porn and rank that first. The other one has no preference for porn and instead ranks the consumption of art first. When abstracting these preferences, most SWFs would result in a ranking of possible social states where the social state in which porn consumption yields the highest total welfare is the best possible state. That is not intuitively appealing. While many of us would give priority to people's preferences even if these preferences are not our own and even if we deem them base, most would stop at evil preferences.⁶

Hence a SWF based on abstract preference satisfaction is not acceptable. Perhaps we should therefore reconsider the Platonism scolded by Arrow⁷ (but without the paternalism). I propose to do just that and more specifically to use human rights as the Platonic good. I have good reasons for choosing this good, and I've defended those [elsewhere](#). You may or may not agree with these

³ The capabilities approach - which argues that freedom (what people are free to do or be) should be included in welfare assessments - is one interesting formulation of well-being, but it's inevitably controversial.

⁴ K. J. Arrow, *Social Choice and Individual Values*; 2012, Yale University Press, p. 23.

⁵ L. Robbins, *An Essay on the Nature and Significance of Economic Science*, 1932, MacMillan, p. 124.

⁶ There are numerous other problems with preference satisfaction as a definition of welfare, e.g. preference manipulation, people having preferences about other people's preferences (as in [Sen's Liberal Paradox](#), etc.).

⁷ See note 4 above. One can argue that welfare economics is always and inherently normative. Disagreement about F is indicative of the ethical choices inherent in SWF thinking. See notes 1 and 2 above. In addition, the focus on individual welfare as the only source of social welfare also marks an ethical choice, namely the choice against collectivism.

reasons, and the reasons will inevitably be controversial. Introducing controversial elements into a SWF may make it useless, and may result in as many SWF as there are people. However, the controversial elements have always been there. The advantage of human rights is that there is widespread agreement on them - at least on the list of rights, not so much on [the reasons why we need them](#).

Another advantage of the use of human rights as a proxy of welfare is interpersonally comparability. Thomas Scanlon, although not talking about human rights, says it better:

Giving up the idea that value judgments can be avoided altogether allows us to make, within moral argument, the kind of interpersonal comparison mentioned earlier. We can still pursue the aim of Neutrality by basing our moral arguments on a conception of the elements that are important in making a life good that is at least widely shared.⁸

Measurement of human rights is also less of a problem compared to alternative measures of welfare. People's rights are respected or not. To some extent, this can be objectively ascertained, at least in theory. There are, of course, practical obstacles to human rights measurement, especially in oppressive environments,⁹ but a reasonable way out of this would be to assume that impossibility of measurement equals absence of respect. Indeed, when rights are not respected, they may be violated in degrees. A SWF based on rights should therefore be able to include measurement of measurable degrees of respect/violation of rights. Again, there is no real obstacle to do this kind of measurement of degrees, at least in the case of most human rights. Free speech for instance is already measured by degrees.¹⁰

In order to convince traditional welfare economists of the usefulness of this approach to SWFs, it's probably best to point to the fact that human rights also encompass a notion of welfare. Traditional welfare needs such as food, shelter, rest, preservation of life, security and a basic standard of living are all codified as human rights.¹¹ Human rights take a [sufficientarian approach](#) to these issues, meaning that the focus should be on the welfare needs of those members of society who do not yet have the means to satisfy them. Improving the standard of living of those who already have an above threshold level is not required by human rights.

This sufficientarianism of human rights suggests that the SWF should be Rawlsian, focusing on the welfare of the least well-off individual member of society. That is true for the welfare element of human rights, but not for the more traditional human rights. The members of society whose rights - welfare and other rights - are most precarious should of course receive priority as to their rights protection, but at least in the case of non-welfare rights this is only an intermediate goal. The SWF should allow us to advance the rights of all members up to a level of equal protection.

[Sen's social welfare function](#) is even less useful in this context because it focuses too much on income and income inequality and doesn't capture the full range of human rights. Completely useless are the Benthamite/classical utilitarian SWF: both the sum and the average are more or less indifferent to the distribution of utility among individuals (with the exception of diminishing marginal utility). Human rights are equal rights, hence their distribution is of crucial importance. Perhaps a Bernoulli-Nash SWF in which F is the product? Maybe, but that will only be the start. A complete SWF based on rights requires a lot more work than I can possibly accomplish here. For

⁸ T.M. Scanlon, *The Moral Basis of Interpersonal Comparisons*, in Elster & Roemer, *Interpersonal Comparisons of Well-Being*, Cambridge University Press, 1993, p. 44.

⁹ See [Human Rights, the Quantitative Approach](#), F. Spagnoli

¹⁰ Freedom House and others are known for this.

¹¹ Although there is opposition to the inclusion of these rights in the corpus of human rights, especially in the U.S. See [here](#).

example, one element of complexity is geography: normally, SWFs are made for a single country, whereas a rights based SWF should of course encompass the whole of humanity. Human rights are rights of all human beings, and there are substantial cross-border effects on rights. I hope to flesh out this proposal another time.