ECON320 Notes II: Lecture Notes on Social Choice Theory Part I

Martin Kaae Jensen

March 23, 2005

Correspondence: Martin Kaae Jensen, Department of Economics, Ashley Building, University of Birmingham, Edgbaston, Birmingham B15 2TT, Birmingham, UK.
E-mail: M.K.Jensen@bham.ac.uk.

Contents

1 Preliminaries

For a set of alternatives X, a *preference relation* $\succeq$, is a (binary) relation which for every two elements from X, x and y, say, either has $x \succeq y$ true or not true. If it is true, that is if $x \succeq y$, we say that "x is preferred to y". The preference relation is *complete* (also called *total* in the literature) if for all $x, y \in X$ we either have $x \succeq y$ *or* $y \succeq x$ *or* both. If both hold, then we say that the agent is *indifferent* between x and y and write $x \sim y$. If x is preferred to y but the agent is *not* indifferent between them, we write $x \succ y$ which reads "x is strictly preferred to y". In addition to completeness, we shall also need $\succeq$ to be *transitive*: $x \succeq y$ and $y \succeq z$ implies $x \succeq z$.

Throughout these notes, preference relations are always assumed to be both complete and transitive which, for short, we express by saying that the preference relation is *rational*. Note that an agent whose preference relation is rational is - because of the completeness part - *always* able to make a choice between two alternatives (the choice possibly being that she is indifferent). So, no matter how complex the question posed, say: "when you are ninety and need medicine, do you want me to give you drug A or B", the agent can form an opinion and so cannot answer "I do not know". This should be kept in mind, because it means of course that our theory may be unrealistic if applied to certain situations.

An alternative way of describing an agents preferences is by means of a utility function $u : X \to \mathbb{R}$ which to each $x \in X$ assigns a level of utility $u(x)$. The utility function u is said to *represent the preference relation* $\succeq$ if for all $x, y \in X$: $u(x) \geq u(y) \Leftrightarrow x \succeq y$. It is an easy matter to show that if a preference relation can be represented by a utility function, then it must be rational. Conversely, one can prove that if $\succeq$ in addition to being rational is continuous[1] and X satisfies certain regularity conditions (convexity will do, or alternatively if X is a finite set we are ok), then there will exist a continuous utility function which represents it.

Theorem 1 *Assume that $u^i : X \to \mathbb{R}$ represents the (rational) preference relation $\succeq^i$. Then $\tilde{u}^i = \phi \circ u^i$ also represents $\succeq^i$ whenever $\phi : \mathbb{R} \to \mathbb{R}$ is a strictly increasing function.*

Which representation (utility or preferences) is more convenient depends upon circumstances, and so we shall use both in what follows.

2 Social Welfare Functions

The point of view in social welfare theory is that of a *social planner* who makes a social choice on behalf on, and affecting, the agents in a society. In practice, the social planner may be a politician in the usual sense, or it may be a society's laws whose implementation are, in principle, just a matter of authority. The description of the social planner we aim for is inspired by the individual consumers in general equilibrium theory: The social planner has a rational preference relation defined on a set of alternatives.

[1] The relation $\succeq$ is continuous if for any two convergent sequences $(x_n)_{n \in \mathcal{N}}$ and $(y_n)_{n \in \mathcal{N}}$, $x_n \succeq y_n$ for all n implies that $\lim_{n\to\infty} x_n \succeq \lim_{n\to\infty} y_n$.

2.1 The set of alternatives

We begin with the *set of alternatives*, X. X could be the set of feasible allocations in an exchange economy for given initial resources:

Example 1 $X = \{x = (x^1, \ldots, x^I) \in \mathbb{R}_+^{NI} : \sum_i x^i = \sum_i e^i\}$. *In the case with two consumers and two goods, X is just all of the points in the Edgeworth box.*

In general we should allow X to be whatever suits the application we have in mind, in particular X does not have to be a subset of $\mathbb{R}^n$ but could consist of elements such as "yes" and "no":

Example 2 *Consider a country with $I = 5.000.000$ citizens ruled by a president who may or may not stay in office. The alternatives for each agent is that the president stays in office ("yes"), or that the president does not stay in office ("no"). The set of alternatives is consequently $X = \{yes, no\}$.*

2.2 Preference relations over the set of alternatives

I agents, which may be consumers, voters, warriors, or what have you, have preferences over the alternatives in X. More precisely, $\succeq^i$ is assumed to be a rational (complete and transitive) preference relation, so in particular either $x \succeq^i y$, or $y \succeq^i x$, or both ($x \sim^i y$), for all $x, y \in X$.

Example 3 *Let agent i face the set of alternatives of example 2, i.e., let $X = \{yes, no\}$. The preference relation $\succeq^i$ then expresses whether the agent prefers "yes" to "no", "no" to "yes", or both (is indifferent). So either yes $\succ^i$ no or no $\succ^i$ yes, or no $\sim^i$ yes.*

Sometimes one has to be a little careful when defining the preference relation because - unlike in the case of a competitive economy - each $\succeq^i$ is defined upon X (and not, for example, on the consumption vector of consumer i). However, if X consists of feasible consumption vectors (cf. example 1), one can always take $x \succeq^i y$ to mean that $x^i \tilde{\succeq}^i y^i$ where $\tilde{\succeq}^i$ is a preference relation in the usual sense of the consumer.[2] Later we shall return to this in more detail.

2.3 Social Welfare Functions

The idea embodied in social welfare functions, is that a social planner is able to form a ranking of the elements in X based upon knowledge of the preference relations of the I agents $(\succeq^1, \succeq^2, \ldots, \succeq^I)$. For example, a social planner might favor the feasible allocations of example 1 to be *equally* distributed among the I consumers. But if this is the case, what are the social planner's preferences and is there a sense in which such a rule may not be a very good, or rational, or just one ? The first of these questions is

[2] Note that $x = (x^1, \ldots, x^I)$ and $y = (y^1, \ldots, y^I)$), so the previous construction simply means that consumer i only cares about his part of the allocation, namely x^i (resp. y^i).

roughly what the social welfare function is supposed to provide a systematic answer to. The second will be dealt with later on.

As mentioned, the social planner's ranking is supposed to be based on the preferences of all agents $(\succeq^1, \succeq^2, \ldots, \succeq^I)$. The sequence $(\succeq^1, \succeq^2, \ldots, \succeq^I)$ is from now on referred to as a *preference profile*. If every preference relation is rational we call this a *rational preference profile*. The set of all possible rational I-agent preference profiles is denoted $\mathcal{R}^I$, so an element in $\mathcal{R}^I$ is a sequence $(\succeq^1, \succeq^2, \ldots, \succeq^I)$ of individually rational preference relations. At first, the notion of a social welfare function may be a bit confusing because it is defined on these preference profiles. Intuitively, one way to think of the situation is by viewing two different preference profiles $(\succeq^1, \ldots, \succeq^I)$ and $(\succeq^1_*, \ldots, \succeq^I_*)$ in $\mathcal{R}^I$, as two different *groups of people*. A social welfare function then must "apply equally to both of these", *i.e.*, the social planner must be able to evaluate social welfare for both groups by some "meta-criterion" which does not take into account the specifics of the groups. In fact, we require in our first definition that the social planner is able to determine a social preference relation for *any* possible preference profile.[3]

Definition 1 (Social Welfare Function) *A social welfare function is a mapping F which to each possible rational preference profile, $(\succeq^1, \succeq^2, \ldots, \succeq^I) \in \mathcal{R}^I$, assigns a rational social preference relation $F(\succeq^1, \succeq^2, \ldots, \succeq^I) \in \mathcal{R}^1$.*

Notice that the actual social preference relation is denoted $F(\succeq^1, \succeq^2, \ldots, \succeq^I)$. That $F(\succeq^1, \succeq^2, \ldots, \succeq^I) \in \mathcal{R}^1$ means that this is a rational 1-agent preference relation. If we write $xF(\succeq^1, \succeq^2, \ldots, \succeq^I)y$, this means that x is *socially preferred* to y. If x is *strictly socially preferred* to y (*i.e.*, $xF(\succeq^1, \succeq^2, \ldots, \succeq^I)y$ but not $yF(\succeq^1, \succeq^2, \ldots, \succeq^I)x$) we write $xF_p(\succeq^1, \succeq^2, \ldots, \succeq^I)y$. Finally, we say that x and y are *socially indifferent*, and we write $xF_I(\succeq^1, \succeq^2, \ldots, \succeq^I)y$, if $xF(\succeq^1, \succeq^2, \ldots, \succeq^I)y$ and $yF(\succeq^1, \succeq^2, \ldots, \succeq^I)x$. An example will help clarify the situation.

Example 4 (Majority Voting) *Consider again the set $X = \{yes, no\}$ (a referendum, "should the president continue in office"). Each of the I voters has three possibilities, he can vote "yes", "no", or choose not to vote (interpreted as "indifference"). If he votes "yes", his preference relation is revealed as having yes $\succ^i$ no. If he votes "no", then no $\succ^i$ yes. If he does not vote then yes $\sim^i$ no. You should convince yourself that from any one of these three possibilities we (or the social planner) actually has a complete description of the agent's preference relation. One social welfare function - and a very important one to be sure - is that of majority voting. This is given by the following $yesF_p(\succeq^1, \ldots, \succeq^I)no$ if and only if "the number of yes votes is greater than the number of no votes", $noF_p(\succeq^1, \ldots, \succeq^I)yes$ if and only if "the number of no votes is greater than the number of yes votes", and finally, $yesF_I(\succeq^1, \succeq^2, \ldots, \succeq^I)no$ if and only if "the number of yes votes is equal to the number of no votes".*[4]

[3]So, we can also think of this as a situation where a social planner must fix a social welfare function *before* actually knowing the agents' preference relations.

[4]In fact, this example is a slightly misleading because it blends the mechanism by which the planner gains knowledge of the agent's preference relations with the actual social welfare function (see the next subsection for discussion). Nevertheless, its simplicity and familiarity makes it an excellent introductory example !

Take notice that in the above definition we require the social preference relation to be rational (complete and transitive). In the majority voting example this is surely the case: Completeness means that any two alternatives can be compared. But there are only two alternatives (!), namely "yes" and "no", and the social planner is of course able to compare these. As far as transitivity is concerned, you may verify that this is actually trivial when there are only two alternatives.

Also note the central aspect which is that the social planner through F is able to evaluate social welfare for *any* possible preference profile of the agents (as long as these are rational, which is also featured as part of the definition). While this is an entirely logical requirement in the situation with a referendum with two alternatives, it is not always so. In fact we shall be *needing* a definition similar to the one above, but taking into account that F may only be defined on a subset of $\mathcal{R}^I$.

Definition 2 (Social Welfare Function with Restricted Domain) *A social welfare function on a subset $\mathcal{A} \subseteq \mathcal{R}^I$ of all rational preference profiles is a mapping F which to each preference profile $(\succeq^1, \succeq^2, \ldots, \succeq^I) \in \mathcal{A}$, assigns a (rational) social preference relation $F(\succeq^1, \succeq^2, \ldots, \succeq^I) \in \mathcal{R}^1$.*

In this definition, the set $\mathcal{A}$ is a subset of all rational preference profiles. An example follows:

Example 5 *Let $\mathcal{A}$ denote the set of preference profiles $(\succeq_1, \ldots, \succeq_I)$ with the property that for every i, $\succeq_i$ can be represented by a utility function of the form $u^i(x) = \sum_{k=1}^n x_k^\sigma$, $0 < \sigma < 1$. Since a preference relation which has a utility representation is always rational (see section 1), $\mathcal{A}$ is indeed a subset of $\mathcal{R}^I$.*

So, in the previous example we single out a subset of all rational preference relations - in fact a very small subset. Demanding that a social planner is able to form preferences only for such a "special" group of consumers is intuitively much less demanding than it is if the planner must "form an opinion" for *any* possible preference profile. As we shall see, this intuition is very sound.

This subsection's last definition is quite natural, and has been implicit in most of the above: Given the social preference relation $F(\succeq^1, \ldots, \succeq^I)$, it singles out the alternative(s) (there may be more than one, though these most be socially indifferent) which according to the social planner should be implemented.

Definition 3 (Social Optimum) *An alternative $x^* \in X$ is socially optimal given the social welfare function F, if $x^* F(\succeq^1, \succeq^2, \ldots, \succeq^I) x$ for all $x \in X$.*

Example 6 *In the majority voting case of example 4, the socially optimal outcome ("yes" or "no") is simply that which gets most votes (if any) because this will then be strictly socially preferred to the other alternative. If none of the two possibilities get a majority of votes so that yes$F_I(\succeq^1, \succeq^2, \ldots, \succeq^I)$no, both "yes" and "no" will be socially optimal in the sense of definition 3.*

With definition 3, we have completed the formal description of what we first stated loosely: A social welfare function is a "rule" by which a planner (concrete or virtual) forms preferences given the individual agents' preferences (the preference profile). Once the social welfare function is "fed" with a concrete preference profile, the social "recommendation" is the social optimum. Whether the social optimum can actually be implemented is, of course, an entirely different question.

2.4 A Look Back

In the following SP stands for Social Planner. Roughly, the picture we have so far drawn is the following:

SP collects information about preference relations $\rightarrow$
SP evaluates social welfare through F $\rightarrow$
SP singles out a socially optimal subset $\rightarrow$
SP - or its police force - implements a choice $x^* \in X$ which is socially optimal given F.

Out of these stages, we have analyzed the second in detail. Each of these stages may not be possible/feasible, but we shall for the most part ignore any such potential difficulties here. Still, a few words seem at their place. First, it may not be possible for a social planner to collect the required information - after all this is potentially a lot of information, and what is more, there is often a sense in which the planner cannot just "go and ask", because agents are then likely to manipulate the planner by misrepresenting their preferences. Now, even if the social planner can gather this information - an example being a vote under ideal conditions - it may not be able to implement the choice x^*. There are cases where it can, though (at least in principle): In the case of a competitive economy, we know from the second welfare theorem that any feasible allocation can be implemented as a Walrasian equilibrium after a suitable redistribution of intial resources. So, if such a redistribution is possible, the socially preferred outcome can actually be realized. Another example is a well specified referendum, where a set of laws are backed up by authority (ultimately force), and so at least in relatively stable societies leads to the implementation of the social optimum.

3 Postulates on Social Welfare Functions

In subsection 2.4, we briefly discussed the stages involved in a social decision process and the problems which may be associated with each stage. One can discuss such things for a long time, and there is a lot of work out there dealing with such issues. The more basic problem - which we should handle first - is whether and under what conditions, the concept of a social welfare function makes any sense in the first place. Indeed, if it does not, who cares whether its optimum can be implemented ?

The way we proceed is by presenting certain postulates/axioms which we think it is reasonable for a social welfare function to satisfy. When faced with a specific social welfare function we can then ask if this is acceptable in the sense of one or all of these postulates. Furthermore, we can list a number of postulates and ask whether there exists a social welfare function at all which satisfies all of them. More on this later.

In the following we take our starting point in definition 2, *i.e.*, social welfare functions defined on a restricted domain $\mathcal{A} \subseteq \mathcal{R}^I$. We allow for the possibly that $\mathcal{A} = \mathcal{R}^I$, in which case we say that the domain is *unrestricted*.

Definition 4 (Paretian) *A social welfare function F is Paretian (or, satisfies the Pareto condition), if for any preference profile $(\succeq^1, \ldots, \succeq^I) \in \mathcal{A} \subseteq \mathcal{R}^I$ and any pair of alternatives $x, y \in X$, x is socially preferred to y whenever every agent prefers x to y, i.e., if for all $x, y \in X$, $x \succeq^i y$, $i = 1, \ldots, I \Rightarrow xF(\succeq^1, \ldots, \succeq^I)y$.*

All social welfare functions which we consider in these notes will be Paretian. In fact, the Pareto condition is a minimum requirement on social welfare functions which we are not prepared to sacrifice. Of course one could easily imagine a social planner who "does whatever he wants" without regarding the preferences of the individual agents. But here we wish to examine social choices in a context where the social planner actually respects the individual agents' tastes. And once we have made *that* decision there is no way around the Pareto condition for if a social welfare function (SWF) did not satisfy it, the planner would actually *overrule* a subchoice which *all* of the agents agree upon.

If we consider two preference relations on X, $\succeq^i$ and $\succeq^i_*$, we say that "$\succeq^i$ and $\succeq^i_*$ rank the elements $x, y \in X$ in the same way" provided that if $x \succeq^i y$ then $x \succeq^i_* y$ and conversely, and also: if $y \succeq^i x$ then $y \succeq^i_* x$ and conversely. Mathematically we can also write this as: (i) $x \succeq^i y \Leftrightarrow x \succeq^i_* y$, and (ii) $y \succeq^i x \Leftrightarrow y \succeq^i_* x$. Perhaps the most intuitive expression reads: either (i) $x \succeq^i y$ and $x \succeq^i_* y$, or (ii) $y \succeq^i x$ and $y \succeq^i_* x$, or (iii) both.

Example 7 *Let $X = \{apples, oranges, bananas\}$ and consider a group consisting of I consumers. Imagine that the group of consumers has preference profile $(\succeq^1, \ldots, \succeq^I)$ in the morning and preference profile $(\succeq^1_*, \ldots, \succeq^I_*)$ in the evening. If consumer $i \in \{1, \ldots, I\}$ prefers apples to oranges in the morning (i.e., apples $\succeq^i$ oranges) if and only if the consumer preferes apples to oranges in the evening (i.e., apples $\succeq^i_*$ oranges), AND if the consumer prefers oranges to apples in the morning (i.e., apples $\preceq^i$ oranges) if and only if the consumer preferes oranges to apples in the evening (i.e., apples $\preceq^i_*$ oranges), THEN $\succeq^i$ and $\succeq^i_*$ rank apples and oranges in the same way. If this is the case for all i, we say that the preference profiles $(\succeq^1, \ldots, \succeq^I)$ and $(\succeq^1_*, \ldots, \succeq^I_*)$ rank apples and oranges the same way. Note - and this is very important to understand - that the statement that apples and oranges are ranked the same way* does not *actually specify whether the agents prefer apples to oranges, or oranges to apples (or both=indifference).*

The next axiom is notoriously difficult to understand if one forgets this notion of "ranking the same way" - so you'd better learn it !

Definition 5 (Independence of Irrelevant Alternatives) *A social welfare function F satisfies the* independence of irrelevant alternatives condition *if for any two preference profiles $(\succeq^1, \ldots, \succeq^I) \in \mathcal{A}$ and $(\succeq^1_*, \ldots, \succeq^I_*) \in \mathcal{A}$ and any two alternatives $x, y \in X$ which each agent i ranks the same way under $\succeq^i$ and $\succeq^i_*$, the social preference relations $F(\succeq^1, \ldots, \succeq^I)$ and $F(\succeq^1_*, \ldots, \succeq^I_*)$ also rank x and y the same way.*

The independence of irrelevant alternatives condition also goes under the name "pairwise independence condition". Whatever one calls it, it is a fact that it is much easier to understand and remember in words than in its formal formulation, so you should spend some time familiarizing yourself with its meaning.

Independence of irrelevant alternatives is, at least from a mathematical perspective, a strong condition. Remember that the Pareto condition is a mimumum requirement on the SWF to the effect that the SWF "respects" individual preferences. Independence of irrelevant alternatives can be interpreted in a similar fashion - though it is no longer a self evident implication of "respect for the individuals". It says that if two groups of agents (or the same group in the morning and the evening as in the example above), $(\succeq^1, \ldots, \succeq^I)$ and $(\succeq^1_*, \ldots, \succeq^I_*)$ were to actually agree on the ranking of two alternatives, then the social preference relations $F(\succeq^1, \ldots, \succeq^I)$ and $F(\succeq^1_*, \ldots, \succeq^I_*)$ should *also* agree on this ranking. If we think of these two groups as actually being the same individuals as in the apples-bananas-oranges example, this sounds quite reasonable. But if we mean, say, people in different countries, it becomes much more difficult to defend the postulate because it involves a measure of interpersonal comparison. Of course, there is also the direct interpretation: That if two alternatives are ranked unambiguously by two preference profiles then the SWF should not depend on any other information than these two alternatives (thus the term "independence of irrelevant alternatives"). But did you actually get this ? Are you sure ? Well, the point is that one should be *very* careful when imposing such "axioms" each of which by itself may sound reasonable enough, for the consequence may be very strong due to some small technical detail which our intuition did not really grasp. The discipline where one makes a number of postulates and accepts each of these *individually*, after which one derives some wild consequence of them *collectively* is a branch of sophism and was very popular at the market square in Ancient Greece (supposedly, Socrates was an expert). So, be careful !

3.1 Arrow's Impossibility Theorem

Arrow's impossibility theorem states, roughly, that if a social welfare function (SWF) with an unrestricted domain (cf. definition 1) is to be both Paretian and satisfy the independence of irrelevant alternatives condition, then this social welfare function must be *dictatorial*. Dictatorial in turn is defined as follows:

Definition 6 (Dictatorial SWF) *A social welfare function F is dictatorial if there is an agent $j \in \{1, \ldots, I\}$, such that for any preference profile $(\succeq^1, \ldots, \succeq^I)$ and any pair of alternatives $x, y \in X$, $x \succ^j y \Rightarrow x F_P(\succeq^1, \ldots, \succeq^I) y$, i.e., if x is strictly socially preferred to y whenever agent j strictly prefers x to y.*

If a SWF is dictatorial for agent j, then we could fix any preference relation for this agent $\succeq^j$ and the SWF would agree with this (in the sense of definition 6) *regardless* of the other agents' preference relations. Moreover, if agent j was to change her preferences from $\succeq^j$ to $\succeq^j_*$, say, the SWF would follow suit. Taken together it is clear now, that a dictatorial SWF establishes a dictatorship by agent j, and so - by the same line of reasoning as was used when describing the Pareto condition above - is unwarranted from a normative perspective.

Theorem 2 (Arrow's Impossibility Theorem) *Assume that there are at least three alternatives in X and that the domain of social welfare functions is unrestricted, i.e., $F : \mathcal{R}^I \to \mathcal{R}^1$. Then any social welfare function which is Paretian and satisfies the independence of irrelevant alternatives condition is dictatorial.*

We will not attempt to prove Arrow's impossibility theorem in this course. There are quite a few different proofs out there, but they are all too time consuming for our quite limited number of lectures ! The presentation given above is close to that of Mas-Collel et al., so if you wish to read the proof you should consult that book first.

3.2 Discussion of Arrow's Impossibility Theorem

What does Arrow's impossibility theorem tell us ?

At a glance it says that it would not be possible for a society to find a social welfare function defined upon *all* preference profiles which is not dictatorial and at the same time satisfies the two normative conditions of definitions 4 and 5.[5]

This is somewhat depressing, although it does not say, of course, that social choices cannot be made. For one thing social choices *are* made in the real world - so at the very least a descriptive approach to social choice should be feasible. Inevitably, such an approach will lead one to study real word social decisions carefully - and one should then try to develop a theory which takes the real word into account, so to speak. This has, of course, been done. As a matter of facts, political science can to some extend be defined as the branch of science which concerns itself with such questions.

A response more in line with these notes, *i.e.*, a more economic (microeconomic, if you like) approach would question the various assumptions (whether technical or axiomatic) behind the result. There is an enormous amount of literature which takes up these challenges and seeks to weaken Arrow's conditions in order to find situations in which social welfare functions *can* be defined. In the next part of this collection of notes, we shall look into such matters.

At the end of the day, Arrow's impossibility theorem teaches us an important lesson: that there is no universal short cut to just and reasonable evaluation of social welfare. In particular this implies that we need to study the situation in more detail, so this is what we will do !

4 Social Welfare Functions Defined on Utility Functions

It is often convenient to assume that $\succeq^i$ can be represented by a utility function $u^i : X \to \mathbb{R}$ (section 1). As explained, we need further assumptions on $\succeq^i$ and X in order to be able to find such a utility representation, but in these notes we shall not worry ourselves with this problem and so simply consider utility functions when it is convenient or the situation demands it.

[5] In addition, there must be at least three alternatives and at an even more technical level, the social welfare function must yield a rational social preference relation.

Recall from section 1 that a utility function u^i represents the preference relation $\succeq^i$ if for all $x, y \in X$: $u^i(x) \geq u^i(y) \Leftrightarrow x \succeq^i y$. If u^i represents $\succeq^i$, then so does any monotone transformation of u^i (theorem 1).

Definition 7 (Social Welfare Function: Utility Formulation) *A social welfare function is a function U which to each possible vector of utility functions for the I agents $(u^1, u^2, \dots, u^I)$ assigns a social utility function $U(u^1, u^2, \dots, u^I) : X \to \mathbb{R}$.*

As in the case with preference relations, we say that x is socially preferred to y if $U(u^1, u^2, \dots, u^I)(x) \geq U(u^1, u^2, \dots, u^I)(y)$, strictly socially preferred if $U(u^1, u^2, \dots, u^I)(x) > U(u^1, u^2, \dots, u^I)(y)$, and that x and y are socially indifferent if $U(u^1, u^2, \dots, u^I)(x) = U(u^1, u^2, \dots, u^I)(y)$.

Example 8 *Consider an exchange economy $(\tilde{u}^i, e^i)_{i=1}^I$ with $n \in \mathbb{N}$ goods. The feasibility set is $X = \{(x_1, \dots, x_I) \in \mathbb{R}_+^{IN} : \sum_i x^i = \sum_i e^i\}$. Let X be the set of alternatives. Given an alternative $x \in X$, agent i's utility is $u^i(x) = \tilde{u}^i(x^i)$.*[6] *As an example of a social welfare function let $U(u^1, \dots, u^I)(x) = \sum_i u^i(x) = \sum_i \tilde{u}^i(x^i)$, $x \in X$. This social welfare function "spits out" a social utility function which weights each of the I consumer's individual utility functions equally.*

The social welfare function in the previous example is one of the most frequently studied in the literature. In fact it is so important that it shall get its own definition !

Definition 8 (Utilitarian Social Welfare Function) *A social welfare function $U = U(u^1, \dots, u^I)$ is utilitarian if it has the form:*

$$U(u^1, \dots, u^I) = \sum_{i=1}^{I} u^i$$

i.e., $U(u^1, \dots, u^I)(x) = \sum_{i=1}^I u^i(x)$.

while we're at it, we may as well introduce another SWF defined upon utility functions. As with the utilitarian, this plays an important role in the literature.

Definition 9 (Rawlsian Social Welfare Function) *A social welfare function $U = U(u^1, \dots, u^I)$ is Rawlsian if it has the form:*

$$U(u^1, \dots, u^I) = \min\{u^1, \dots, u^I\}$$

i.e., $U(u^1, \dots, u^I)(x) = \min\{u^1(x), \dots, u^I(x)\}$.

Intuitively, the Rawlsian form view's the society's objective as that of maximizing the utility of the individual who is *least* well off. Sometimes, this criterion is referred to as maximin (or a variation hereupon such as max-min or maxmin). At any rate, the Rawlsian criterion should be viewed as an expression of social *justice* according to Rawls

[6] Notice here how we go from individual utilities $\tilde{u}^i$ which are defined upon agent i's consumption set $\mathbb{R}_+^n$ to individual utilities u^i which are defined upon the set of alternatives $X \subset \mathbb{R}_+^{IN}$.

(and one is entitled, of course, to agree or disagree with such a criterion).

It is crucial that you understand the potential discrepancy between the notions of a social welfare function defined in terms of utility functions and preference relations. Take the utilitarian social welfare function. Here, given a vector of utility functions, the resulting social utility function $U(u^1, \ldots, u^I)(x)$, will come to depend on the actual representation chosen for each of the agents. Thus if, say $u^1(x)$ is replaced with $100u^1(x)$ (which represents the same preference relation!), the social utility function will be different: it will place hundred times more weight on consumer 1. Obviously, this makes absolutely no sense if we interpret utility functions in the way we usually do in general equilibrium theory: namely as *ordinal* measures. On the other hand, we could depart from this ordinal approach in favor of an approach where the actual utility level $u(x)$, and not just the ranking, is of significance. Such a utility concept is called *cardinal* and so we have seen now two examples of social welfare functions defined upon utility functions which *must* be viewed from such a cardinal perspective to have any normative meaning. The appeal in this cardinal approach is obvious: One is able to make *interpersonal* comparisons such as that which is featured in the expression "the individual who is least well off". With the purely ordinal approach such a statement makes no sense: We can say that x is preferred to y for agent i, but we cannot judge whether, say, i prefers x to y *more* than another agent does.

How then, does on "recover" the ordinal nature of preferences as expressed in the preference formulation. The answer is given in the following:

Definition 10 (Ordinal SWF) *A social welfare function U is ordinal if for any choice of I strictly increasing transformations $(\phi_1, \ldots, \phi_I)$, $\phi_i : \mathbb{R} \to \mathbb{R}$: $U(u^1, \ldots, u^I)(x) \geq U(u^1, \ldots, u^I)(y) \Leftrightarrow U(\phi_1(u^1), \ldots, \phi_I(u^I))(x) \geq U(\phi_1(u^1), \ldots, \phi_I(u^I))(y)$ for all $x, y \in X$.*

Exercise 4.0.1 *Read the previous definition carefully and convince yourself that it does make the utility and preference formulations of social welfare functions equivalent.*

As explained above, the utilitarian welfare function is not ordinal (we say then that it is cardinal). Likewise, as you should check for yourself, the Rawlsian social welfare function is cardinal. It is clear that for an ordinal social welfare function all of the concepts of the previous section can be defined by simply using that if U is ordinal then there is a one-to-one correspondence with a preference relation.

One condition which it is convenient to consider explicitly is the Pareto property:

Definition 11 (Paretian - Utility Formulation) *A social welfare function U defined on utility functions is Paretian (or, satisfies the Pareto condition), if for any vector of utility functions $(u^1, \ldots, u^I)$ and any pair of alternatives $x, y \in X$; $u^i(x) \geq u^i(y)$ for all i implies that $U(u^1, \ldots, u^I)(x) \geq U(u^1, \ldots, u^I)(y)$.*

As you will hopefully agree, this definition is perfectly natural. Moreover, because there is no element of interpersonal comparison involved in it, it applies equally well to any type of social welfare functions defined on utility functions (put differently: It makes sense whether U is ordinal or cardinal).

Example 9 *The utilitarian SWF is Paretian. Indeed, if for every agent i, it holds that $u^i(x) \geq u^i(y)$, $x, y \in X$, then taking the sum, $\sum_i u^i(x) \geq \sum_i u^i(y)$. But this means exactly that $U(u^1, \ldots, u^I)(x) \geq U(u^1, \ldots, u^I)(y)$ where U is the utilitarian form.*

Example 10 *The Rawlsian SWF is Paretian. Indeed, if for every i, it holds that $u^i(x) \geq u^i(y)$, $x, y \in X$, then the minimum among the "left-hand-sides" must also be larger than or equal to the minimum among the "right-hand-sides". Precisely put $\min\{u^1(x), \ldots, u^I(x)\} \geq \min\{u^1(y), \ldots, u^I(y)\}$, which is the conclusion we were looking for.*