

# Interpersonal Preference Comparison

Enhua Hu

## Introduction

This paper discusses a way of formalizing our intuitions concerning interpersonal preference comparisons for pairs of agents. The first section discusses previous literature in the field. The second section presents the theory informally while justifying its philosophical foundations. The third section shows that the theory can be formalized as an ordering which is complete and transitive given some restrictions. The fourth section shows that the ordering can give rise to a Pareto-optimal social choice function with advantages over traditional ones.

A few concepts need clarification. The paper works within the standard framework of microeconomics, which studies individual behavior and preference. States are represented as vectors  $(x_1, x_2, x_3, \dots)$  of  $\mathbb{R}^n$ , in which each number of the vectors indicates the quantity of a particular good. Each agent has a preference ordering (denoted by  $\succeq$ ) over the states. Rational choice theory further assumes that the preference ordering is complete ( $x \succeq y$  or  $y \succeq x$ ) and transitive (if  $x \succeq y$  and  $y \succeq z$  then  $x \succeq z$ ). Lastly, most preference ordering can be represented by utility functions; this is particularly interesting if the utility functions are continuous.

## I. Survey of Literature

Interpersonal comparison (IC) statements have the form, “A prefers x more than B prefers y”. Many economists believe that an IC statement cannot have a truth value because it uses an ordinal concept of utility, which ranks the preferences but does not capture the intensity of the preferences. Thus, if there is no way of comparing intensity for even single agents, it seems nonsensical to compare intensity between agents.<sup>1</sup>

---

<sup>1</sup> D. M. Hausman, “The Impossibility of Interpersonal Utility Comparisons,” *Mind*. 104:473–490, 1995.

Further, the utility of goods is subjective; there is no objective measurement we can use between agents.<sup>2</sup> More recently, Rossi has proposed an epistemological argument against the possibility of ICs.<sup>3</sup>

Despite this view, I argue that the possibility of IC is real. In addition to the theoretical benefits,<sup>4</sup> we have a common intuition that some ICs are true; we know that a starving child gets more utility from bread than does an adult who is full. If the possibility is denied, then statements like the above have no truth value, which is highly counter-intuitive.

The section will survey two main theories for IC and point out their issues. They are:

- Extended sympathy, advocated by Harsanyi<sup>5</sup> and others.<sup>2</sup>
- Zero-one rule, introduced by Isbell;<sup>6</sup> Hausman<sup>1</sup> is a proponent.

### *Extended Sympathy/Empathetic Preference*

Arrow<sup>5</sup> and Sen<sup>6</sup> call this theory ‘extended sympathy’, Harsanyi calls it ‘empathetic preference’, and it is by far the most studied theory of IC. The assertion of empathetic preference is that individuals can empathize with others. Theorists agree that there are generally two processes that must be accomplished for an individual to have a judgment of extended sympathy. First, the individual must undergo an objective shift where she considers herself in the material condition of another agent. Second, the individual undergoes a subjective shift where she replaces her preferences with those of the other agent. These two shifts allow an agent to capture the preference of another via empathetic understanding.

In the following, I will present Harsanyi’s version of the concept. Harsanyi’s model has two axioms:

---

<sup>2</sup> L. Robbins, “Inter-personal Comparisons of Utility: A Comment,” *Economic Journal*, 48:635–641, 1938.

<sup>3</sup> M. Rossi, “Interpersonal Comparison of Utility. The Epistemological Problem,” (PhD Thesis, London School of Economics and Political Science, 2009).

<sup>4</sup> K. Arrow, “Extended sympathy and the possibility of social choice,” *Philosophia*, 7:233–237, 1978.

<sup>5</sup> J. Harsanyi, “Cardinal welfare, individualistic ethics, and the interpersonal comparison of utility,” *Journal of Political Economy*, 63:309–321, 1955.

<sup>6</sup> J. R. Isbell, “Absolute Games,” in A. W. Tucker and R. D. Luce (eds) *Contribution to the Theory of Games, Vol. IV*, Princeton University Press, 357–96.

1. Utility functions, empathetic and personal, satisfy all Neumann-von Morgenstern postulates, such as completeness of preferences.
2. Agents are able to empathize fully with one another and they are able to capture the exact same preference relation. In technical terms, an agent  $i$ 's empathetic function for agent  $j$  is a strictly increasing affine transformation of agent  $j$ 's own preference function.

Harsanyi allows for an agent to maximize the sum of all her empathetic functions by evaluating how much the agent values the utility of one over another agent, a form of IC. Each empathetic function is subjective; therefore, this agent cannot be an impartial social planner as she is. Harsanyi deals with this by putting this agent in a situation of uncertainty such that she must make her decision as if she has equal probability to be any other agent. The idea is similar to Rawls'<sup>11</sup> veil of ignorance, only that Harsanyi makes the agent maximize expected utility rather than follow the maximin principle.

Each empathetic function is subjective and needs to be so since it is one's own IC. Therefore, different agents will likely make different choices when put in this situation. Harsanyi is aware of this; his argument is that agents with enough information will have the same empathetic functions. Second, the idealizations required by Harsanyi are very demanding. In particular, the ability to empathize perfectly and to be an impartial observer may only be conceptually possible.

### *Zero-One Rule*

The zero-one rule was first introduced by Isbell<sup>7</sup> as a criterion of fairness. Unlike empathetic preference, this rule employs utility as cardinal and bounded. The idea is to normalize an agent's utility function such that the utility of the most preferred state is 1 and the utility of the least preferred state is 0.

The theory says that agent  $i$  is better off in state  $x$  than agent  $j$  in state  $y$ , iff the following condition holds:

$$\frac{\text{Max}_i U_i(x) - \text{Min} U_i}{\text{Max} U_i - \text{Min} U_i} > \frac{\text{Max} U_j - U_j(y)}{\text{Max} U_j - \text{Min} U_j}$$

The above can be interpreted as one agent being “closer” to her best state than another. This interpretation only makes sense with cardinal utility because it requires utility to represent intensity. Furthermore it implies that there are most/least preferred states. While these requirements are demanding, the main problem is the hidden assumption that agents have equal capacity for preference. The theory implies that agent  $i$  and  $j$  are equally well off in their worst and best states.<sup>1</sup> Sen argues that a social welfare resulting from this would emphasize society’s resources with lower satisfaction requirements. Given limited goods, the social planner wanting to maximize utility will have to allocate to those who are satisfied with less.

The response from the zero-one rule proponent is to argue that preference satisfaction is what the theory compares, rather than well-being. They argue that well-being is not being compared; rather, the theory compares preference satisfaction. Hausman states:

“No sense has been given to comparing Jill’s non-comparative well-being to Ira’s non-comparative well-being. In the case of cardinal and bounded utilities, the conclusion ought to be that a view of well-being as preference satisfaction leaves interpersonal comparisons undefined and mysterious.”<sup>3</sup>

The argument is that the normalized utility function only describes how well preferences are satisfied and that preference satisfaction is not well-being. The theory evaluates statements of the following form: “Individual  $i$ ’s preferences in state  $x$  are better satisfied than individual  $j$ ’s preferences in state  $y$ .”

If one takes the preference satisfaction approach, then this theory seems more reasonable, albeit less powerful because it has little implication for social welfare. If all we can compare is the extent of preference satisfaction and not the strength of satisfaction itself, then we cannot maximize social welfare. The zero-one rule still leaves a hole to be filled, namely that there is no way of comparing welfare between individuals such that we can assign truth value to statements of the form, “individual  $i$  is better off in state  $x$  than individual  $j$ ”.

## II. Philosophical Issues and Informal Layout

This section lays out informally my theory of IC statements. This exposition captures some of the justifications and intuitions for the theory; I address the more technical and formal concerns in the next section.

The two main problems that face ICs are the lack of objective basis for comparison and the ordinality of utilities. The extended sympathy method is flawed because the objective basis provided is weak. The zero-one rule is flawed because it requires the cardinality<sup>7</sup> of utilities, which goes against standard economics.

My theory deals with the second problem by comparing only preference orders, so the theory always uses ordinal utility. The first problem is more intricate; it might very well be that no objective basis can be found for IC, so my theory proposes intersubjective agreements between agents as the basis for IC. More precisely, statements of the form, “Agent A prefers x more than agent B prefers y” are true if “A thinks she prefers x more than B prefers y” and “B thinks she prefers y less than A prefers x” are true. The claim is that agreements between subjective judgments can entail the truth-bearing quality of IC statements.

### *What is preference?*

IC statements have the form “A prefers x more than B prefers y”; they have two objects and one binary connective. The objects are the preferences the agents have for a good. In standard economics, preference is defined as a relative ranking over the set of goods. One approach would be to define preference for a particular good as the ranking of the good, i.e., the fifth most preferred good. However, this approach fails because many rankings have infinite goods and no most preferred good.<sup>8</sup>

I propose that an agent’s preference for a good is described as the set of all goods to which the agent is indifferent, their ‘indifference set’ for

---

<sup>7</sup> In economics, utility is conceived as either cardinal or ordinal. If cardinal, then the number represents intensity or degree of preference, namely if an apple gives five utility, a pear ten, and a banana three, then we can say that the pear is preferred to the apple *more* than the apple to the banana. If ordinal, then the numbers simply denote the ranking or order of preference. Therefore, all that can be said is that the pear is preferred to the apple which is preferred to the banana.

<sup>8</sup> For instance, a ranking for quantities of a continuous good like water.

that good. When we ask someone how much they value a good, their preference, we ask them to provide an equivalent good. However, we cannot describe an agent's preference for a good,  $x$ , with just one equivalent good,  $y$ . It may be the case that another agent's preference with respect to the two goods ( $x,y$ ) are the same but differ for a third good,  $z$ . If preference is described with respect to only one equivalent good, then the two agents would have the same preference for  $x$  with respect to  $y$  but not with respect to  $z$ . However, preference for a good cannot be subject to other goods but only agents, so we must look at the whole indifference set rather than any particular member when defining preference.

### *Subjective Judgments*

My proposed method uses intersubjective agreements to analyze statements of IC. Subjective judgments have the form, "A considers her preference for  $x$  to be stronger than B's preference for  $y$ ." Once again, this is a binary connective with two objects. The two objects are first, A's indifference set for  $x$ , and second, B's indifference set for  $y$ .

Since this is A's subjective judgment, this binary connective should be A's preference ordering. A's preference ordering allows her to compare individual goods, but gives her no way of comparing sets of goods. I propose that preference over sets of goods be defined as a range, this range defined by the most and least preferred good in that set.<sup>9</sup> Let me clarify this with the following example.

Imagine an agent who is asked to compare her preference for a good,  $x$ , and another agent's preference for a good,  $y$ . She will look at the members of the other agent's indifference set for  $y$ . If it is the case that she considers all those goods superior to  $x$ , then she concludes that the other agent values  $y$  more than she values  $x$ . If our agent considers all those goods inferior to  $x$ , then she concludes that the other agent values  $y$  less than she values  $x$ .

Another case is one in which there are goods in the other agent's indifference set that our agent considers more and less valuable than  $x$ . This case has our agent believing her preference to be vaguely similar to

---

<sup>9</sup> This is slightly technical; I elaborate this further in section III.

her counterpart's. This type of scenario, while unappealing at first, is intuitive and arises often. For instance, two children who like certain chocolates are asked who likes them more.

One might ask, "Why would an agent ever think that her preference is weaker than that of another?" One may think that the agent is not empathetic, or that she could report falsely. A lack of empathy is not a concern because subjective judgments do not require empathy. For false reporting, while most social choice theorists assume to know the true preference of agents, it may not be the case. This is also known as the "Preference Revelation Problem," and there are mechanisms designed so that the agent is forced to report her true preference as it is always her optimal action.<sup>10</sup> This paper will not consider the problem of preference revelation, as it is beyond its scope; instead, I will take the orthodox view that agent preferences are available.

### *Intersubjective Agreement*

Now that I have presented subjective judgments, I will argue that the truth of an IC statement depends on the subjective judgments of the agents whom the IC statement concerns. In particular, if two agents' subjective judgments agree with each other, then the IC statement is true.

This claim is hard to verify because there are no standards for the truth of IC statements. In the following, I propose two views: Either IC statements are equivalent to intersubjective agreements, or the weaker alternative, the truth of intersubjective agreements implies the truth of IC statements.

Recall the problem of objective basis. I will show that in the absence of an objective basis, intersubjective agreements are equivalent to ICs. If there is an objective basis, then intersubjective agreements merely imply IC statements.

Suppose there is no objective basis for IC statements. Then, for the sake of social choice, we must still decide how to best assign truth values for those statements. If it is the case that statements of IC are subjective, then whose subjectivity matters? Clearly, the agents whom the

---

<sup>10</sup> The most famous one being the Vickrey-Clark-Groves Auction.

IC statement concerns have priority. With the absence of an objective ground, the best judgment is one that both agents agree to. If the two agents agree that one has a stronger preference, then an outsider's judgment should not matter. Of course, the two agents may not always agree since not all IC judgments need to be true.

Suppose that there is an objective basis for IC statements, such that there are objectively true and false IC statements. Then individual agents, given enough information, can arrive at the right conclusion regarding ICs. Some statements of IC have seemingly immediate truth interpretations, for instance, "A starving person's preference for food is stronger than a satiated person's preference for food." Other IC statements require more information. Clearly, the agents whom the IC statements concern have the most information, as they know their preference orderings best. Therefore, if the two agents' judgments agree, it is the best approximation of the objective truth. This is similar to what Harsanyi claims, except my method does not require an omniscient social planner.

If one accepts my arguments above, it follows that intersubjectivity allows the analysis of IC statements. One might still ask whether there is an objective basis for ICs. We are now in a position to give some insight into that question. Since, in the absence of an objective basis, IC statements are equivalent to intersubjective agreements, if there are IC statements which are true but do not obtain intersubjective agreements, there must be an objective basis. Unfortunately, we cannot answer this question any further. Since we do not have a formal definition of truth for IC independent of intersubjectivity, we must rely on intuition. Therefore, the set of IC statements which are true will be the intuitively true and obvious ones which will likely always obtain intersubjective agreements. While there may be non-obvious but true IC statements which do not obtain intersubjective agreement, my method cannot analyze them.

### **III. Formal Language and Interpersonal Preference Order**

In this next section I present a logic capable of expressing preference orderings of pairs of agents and I show that we can build an ordering of interpersonal preference from it. Further, this ordering is complete and transitive if the single agent preference satisfies some basic properties.



*Syntax*

The syntax of two agent preference logic is the following:

- The usual logical symbols of predicate logic.
- A set of goods/states:  $S = \{x, y, z, \dots\}$
- A set of binary relations over  $S$ ,  $P = \{\lesssim_1, \sim_1, <_1, \lesssim_2, \sim_2, <_2\}$
- A set of unary relations over  $S$ ,  $Q = \{Q_1^1, Q_1^2, Q_1^3, \dots, Q_n^n\}$

A well-formed formula (wff) is defined the same way as in predicate logic.

*Semantics*

The set of states consists of bundles of goods in  $R^n$ , so each unary relation  $Q_i^j$  denotes that a good has  $j$  quantity of the  $i^{\text{th}}$  component.

Example:  $Q_2^4x$ , denotes that  $x$  has 4 of the 2nd component.

The binary relations capture the preference ordering of two agents, 1 and 2, over the set of states, their meanings are strict/weak preference and indifference.

Example:

$x \sim_1 y$ , denotes that agent 1 is indifferent between  $x$  and  $y$   
 $\forall x(\forall n(Q_n^0x \rightarrow \forall y(x \neq y \rightarrow x <_1 y)))$ , denotes that if  $x$  is an empty state (zero in all components), then any good  $y$  is strictly preferred over  $x$  by agent 1.

In standard economics, it is assumed that all three relations are transitive. However,  $<$  and  $\sim$  are not complete. Only  $\lesssim$  is complete, transitive, reflexive and symmetric. Lastly,  $\sim$  is an equivalence relation while  $<$  is a strict total ordering. Their relationship is as follows:

$$(x \sim y) \leftrightarrow (x \lesssim y \wedge y \lesssim x)$$

$$(x < y) \leftrightarrow (x \lesssim y \wedge \neg (y \lesssim x))$$

*Some definitions*

I now move on to the construction of the interpersonal ordering. First, I define the concept of indifference sets for the evaluation of agent preference over a good. Second, I define interval sets, which captures the concept of subjective judgments.

Def.1 Indifference Set:

Let  $x$  and  $i$  be respectively a bundle of goods and an agent. Denote  $[x]_i$  to be the indifference set of  $x$  by agent  $i$  such that:

$$[x]_i = \{y \mid x \sim_i y \wedge x \neq y\}$$

The indifference set of a bundle of goods,  $x$ , is the set of all goods the agent considers equivalent to  $x$ , excluding  $x$  itself.

The indifference set here is different from standard definitions because I exclude the original good. It would be circular if I defined an agent's preference for a good via use of that good. Further, excluding the original good will allow for more equivalence classes in our interpersonal ordering, making it stricter. However, in the context of continuous utility functions, this assumption is not necessary.

Def.2 Interval Sets

Interval sets can be understood as ways for one agent to evaluate the preferences of another agent.

We say that  $[y, z]_j$  is the interval set for agent  $j$  on indifference set  $[x]_i$  iff

$$\text{For } \forall x \in [x]_i, \forall y \in [y]_j, \forall z \in [z]_j, y \precsim_j x \precsim_j z, \text{ for agent } j$$

Where the indifference sets  $[y]_j$  and  $[z]_j$  contains respectively the least and most preferred goods in  $[x]_i$  by agent  $j$ . The interval set is interpreted as the subjective evaluation of the preference of one agent by another.

Theorem 1

For any non-empty indifference set, there exists one interval equivalence for each other agent.

The proof is shown in the construction of interval sets. Take the most and least preferred outcome in the indifference set by the other agent and assign them as boundaries of the interval. We can do that because the set is not limited by constraints but rather denote possible states.

We have now all the tools needed to define subjective judgments of the form “A considers his preference for x to be stronger than B’s for y”.

Def 3. Subjective Judgments of Preference

Agent  $j$  considers his preference for good  $w$  stronger than that of agent  $i$  for good  $x$ , denoted by  $[x]_i \prec_j [w]_j$ , iff:

$$\forall x \in [y, z]_j, \text{ we have } w \prec_j x,$$

where  $[y, z]_j$  is the interval set of  $[x]_i$  for agent  $j$

In other words, she prefers  $w$  to all goods in her interval equivalence bundle for  $[x]_i$ . An illustration helps one grasp the concept and will be useful as we develop it further.

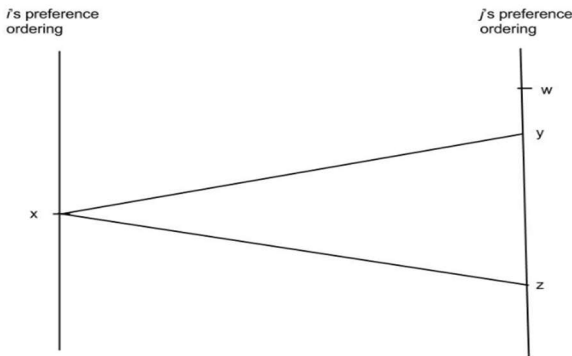


Fig.01:  $[x]_i$ 's interval equivalence  $[y, z]_j$

We can analyze intersubjectivity now that we have defined subjective judgments. There are generally three cases of interval comparison available. The respective intervals can be exclusive from the goods and allow for either intersubjective agreement or disagreement, or the intervals are inclusive of the goods. We can easily interpret the first two cases of exclusivity. If there is intersubjective agreement then we have a strict stronger/weaker preference relation. If there is intersubjective disagreement then the social planner's job is simple since one agent wants what the other one does not want. Graphically, for the two cases we have:

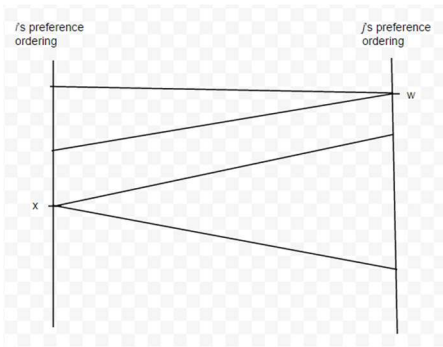


Fig.02 Intersubjective Agreement

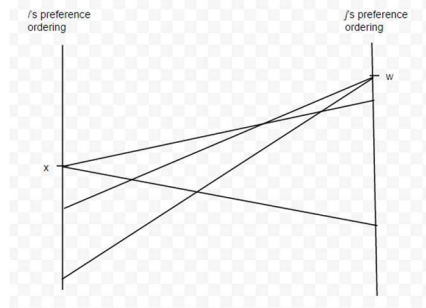


Fig.03 Intersubjective Disagreement

I interpret the case of intersubjective agreement (Fig.02) as “Agent  $i$  prefers  $x$  less than agent  $j$  prefers  $w$ .” For the case of intersubjective disagreement, no interpersonal comparison can be made. Therefore, the interesting cases for the framework exclude those in Fig.03. Fortunately, two mild restrictions on individual agent preference orders exclude the case of intersubjective disagreement.

Def. 4 Strong Monotonicity<sup>11</sup>

Strong monotonicity is a standard axiom in economics; it is defined as follows:

Let  $x$  and  $y$  be bundles of  $n$  goods represented as  $(x_1, \dots, x_n)$  and  $(y_1, \dots, y_n)$ . We say that an agent's preference is *strongly monotonic* iff the following holds:

If  $\exists x_i, y_i$  such that  $x_i > y_i$  and  $\forall x_i, y_i$  are such that  $x_i \geq y_i$  then  $x$  is preferred to  $y$ .

The strong monotonicity property implies that our bundle of goods has, well, goods. More formally, it means that each unit of a good has positive value for the agent.

Def. 5 Income in Bundle

For any bundle of goods,  $x$ ,  $\exists a \in \mathfrak{R}$  such that  $y = (a, 0, 0, \dots, 0)$  and  $y \sim_i x$  for all agents  $i$ .

This just means that our agent can equivocate any bundle with a bundle containing a certain quantity of a single good of the first component. Think of this first component as income; it is not so unreasonable to say that our agent is indifferent between five dollars and five dollars worth of coffee.

Theorem 2

If the two agent preference ordering for  $\preceq$  is transitive and complete while also satisfying monotonicity and income in bundle, then the case of intersubjective disagreement cannot occur. (Proof: *See Appendix.*)

We rule the above case out formally using the two properties for the sake of rigor. Without the two, our system can still make interpretations and is still useful. Furthermore, we could have simply, by assumption, limited our system to analyze preference orderings which do

---

<sup>11</sup> The axiom of strong monotonicity is often replaced by the weaker but more general axiom of local non-satiation. Depending on the context, the two can be equally general and some authors argue that strong monotonicity is implied by economic theory and need not come as an axiom. Below are some papers for the interested reader:  
Becker, G. S. *Economic Theory*, Transaction Publishers, New Brunswick, N.J., 2008.  
Border, K.C. *Lecture Notes: Monotonicity and Local Non-Satiation*, April 2009.

not produce case 2. One may argue that the two properties are too restrictive or unreasonable, but they do not eliminate any real case of interest.

The first property is standard in economics because decisions are usually centered on “good” goods rather than bad “goods.” Furthermore, one can make tweaks to compare preferences over bad goods while respecting this property.

The second property is named “income in bundle;” roughly it assumes the existence of a currency for which all goods can be traded. This need not be money; on a desert island, this might be food or whatever everyone finds valuable. One might argue that food is not currency, but on a desert island, food is potentially more tradeable; there are things people would not trade for money off the island that people would trade for food on a desert island. In short, this property denotes a “prime” good which exists to some extent in all societies. The extent of its tradeability is denoted by what other goods can be traded for it.

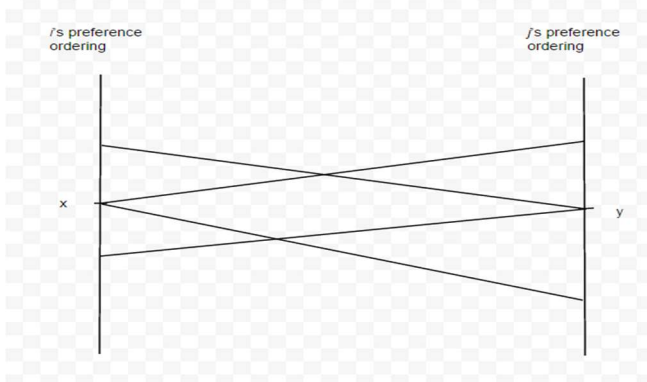


Fig.04 Similar but non-identical preferences (Case 3)

Case 3 is analogous to indifference. This is a case in which the agents cannot intersubjectively agree with each other. We need not rule it out as we did for Case 2 because there is no disagreement in a strict sense.

We now can move on to define the binary connectives of ICs.

The Binary Connectives

We use the same notation for the connectives as the single agent ones for sake of simplicity. It is clear which is which as the single agent ones have subscripts.

Weak Preference Difference

We say that agent  $i$ 's preference for bundle  $x$  is weakly stronger than agent  $j$ 's preference for bundle  $y$  if:

- $j$  considers his preference for  $y$  to be weaker than  $i$ 's preference for  $x$

*or*

- $i$  considers his preference for  $x$  to be stronger than  $j$ 's preference for  $y$

*or*

- $i$  and  $j$  cannot come to agreement, as depicted in case 3.

We denote this by  $[x]_i \succeq [y]_j$ .

Strict Preference Difference

If the two first conditions are satisfied then it is a strict preference difference which is exactly the case of intersubjective agreement. Just like single person preference, the weak preference difference includes a possibility for the strict one.

We denote this by  $[x]_i \succ [y]_j$

Incommensurable

We say that  $[x]_i \sim [y]_j$  if neither agent believes the other's preferences are better satisfied.

Theorem 3

The order produced by  $\succeq$  is transitive and complete.

*Complete:*  $[x]_i \succeq [y]_j$  or  $[y]_j \succeq [x]_i$  must be true.

*Transitive:* if  $[x]_i \succeq [y]_j$  and  $[y]_j \succeq [w]_k$  then  $[x]_i \succeq [w]_k$ .

Proof:

-Completeness: See Appendix.

-Transitivity: Done by breaking the definitions into different cases which all satisfy transitivity.

### Corollary

The three interpersonal relations have the same interpretation as the single agent ones if the two agents being compared are identical.

Namely  $[x]_i \lesssim [y]_i$  is equivalent to  $x \lesssim_i y$ ; the same holds for  $\sim$  and  $\prec$ .

## IV. Social Choice and Welfare

In the following, I will define a social choice function (SCF) using the IC ordering obtained prior. I will then show that it has advantages over the utilitarian and egalitarian social choice functions.

For this section, we will be working with preference ordering which can be represented by continuous utility functions. An SCF is a decision rule. It is a means of choosing a social distribution given the different agents' preferences and a feasibility constraint. A utilitarian SCF is one which maximizes the sum of utilities of the agents; an egalitarian SCF is one that distributes the good equally amongst agents. They are defined formally as:

Utilitarian:  $F(u_i, u_j, z) = (x, y)$  such that  $\max \{u_i(y) + u_j(x) \mid x+y=z\}$

Egalitarian:  $F(u_i, u_j, z) = (x, x)$  such that  $2x=z$

The social choice I propose is the following:

$F(u_i, u_j, z) = \max \{u_i(x) + u_j(y) \mid x + y = z \text{ and } [x]_i \sim [y]_j\}$

This implies that we maximize the sum of utilities, weighting agents equally, as long as neither believes the other agent is better off. We can show that this function is Pareto optimal and produces more total utility than the egalitarian choice, while obtaining a more equal distribution than the utilitarian function.



Theorem 4

The social choice function above is Pareto-Optimal. Namely, the constraint is non-binding with respect to optimality. (Proof: *See Appendix.*)

Theorem 5

The social choice function above always produces more or equal total utility than the egalitarian one. (Proof: Trivial, since the egalitarian distribution always satisfies the additional constraint.)

Theorem 6

The social choice function above always produces a distribution that is more or equally egalitarian than the utilitarian one. (Proof: Suppose that the two distributions are not the same. Then it must be that the utilitarian solution does not satisfy  $[x]_i \sim [y]_j$ . However, this is a constraint on how much allocations can differ, so it must be that the utilitarian solution is allocating goods in a less egalitarian manner. For some cases,  $[x]_i \sim [y]_j$  is not an effective constraint, then the solutions will be the same for the two distributions.)

**V. Appendix****Proof of Impossibility of Case 2 (by contradiction)**

1. Let agent  $i$  strictly prefer his indifference set for  $x$  over the interval equivalence set of  $y$  by agent  $j$ .
2. Let agent  $j$  strictly prefer his indifference set for  $y$  over the interval equivalence set of  $x$  by agent  $i$ .
3. By 1), there is a bundle  $(a, 0, \dots, 0) \sim_i x$  for all the  $x$  in  $[x]_i$ . By strong monotonicity, for all bundles of the form  $(b, 0, \dots, 0)$  in the interval set of  $y$  we have  $a > b$ .
4. By 2), there is a bundle  $(b, 0, \dots, 0) \sim_j y$  for all the  $y$  in  $[y]_j$ . By strong monotonicity, for all bundles of the form  $(a, 0, \dots, 0)$  in the interval set of  $x$  we have  $b > a$ .

**Proof of Completeness (by contradiction)**

Want to show:  $[x]_i \succeq [y]_j$  or  $[y]_j \succeq [x]_i$

1. Let  $\sim([x]_i \succeq [y]_j)$  and  $\sim([y]_j \succeq [x]_i)$
2. Then by  $\sim([x]_i \succeq [y]_j)$ :
  - a)  $j$  does not consider his preference for  $y$  to be weaker than  $i$ 's preference for  $x$
  - b)  $i$  does not consider his preference for  $x$  to be stronger than  $j$ 's preference for  $y$
  - c) Not the case that indifference/Case 3 occurs.
3. And by  $\sim([y]_j \succeq [x]_i)$ :
  - a)  $j$  does not consider his preference for  $y$  to be stronger than  $i$ 's preference for  $x$
  - b)  $i$  does not consider his preference for  $x$  to be weaker than  $j$ 's preference for  $y$
  - c) Not the case that indifference/Case 3 occurs.
4. We see that 2.a)b) and 3.a)b) give us exactly Case 3 which we have assumed would not occur.

### Sketch of Proof of Pareto Optimality

The proof proceeds as follows:

1. I show that the constraint is equivalent to the indifference curves intersecting.
2. I show that the set of allocation where indifference curves intersect contains  $(0,0)$
3. I show that the set of allocation where indifference curves intersect is unbounded.
4. I derive the Pareto Frontier and shows it intersects with the set.
5. Thus, there must be a Pareto solution within the constraint.

**Works Cited**

- Arrow, Kenneth J. "Extended Sympathy and the Possibility of Social Choice." *Philosophia*, 7:233–237, 1978.
- Harsanyi, John. "Cardinal Welfare, Individualistic Ethics, and the Interpersonal Comparison of Utility." *Journal of Political Economy*, 63:309–321, 1955.
- Hausman, Daniel M. "The Impossibility of Interpersonal Utility Comparisons." *Mind*. 104:473–490, 1995.
- Isbell, John R. "Absolute Games." In A. W. Tucker and R. D. Luce (eds), *Contribution to the Theory of Games, Vol. IV*. Princeton University Press, 357–96.
- Robbins, Lionel "Inter-personal Comparisons of Utility: A Comment," *Economic Journal*, 48:635–641, 1938.
- Rossi, Mauro. "Interpersonal Comparison of Utility: The Epistemological Problem." PhD Thesis, London School of Economics and Political Science, 2009.

I would like to extend my thanks to the helpful editors of *Noësis*, in particular Mathew Armstrong and Howard Williams, whose work and feedback have greatly improved this paper. As well, Professor Peski for his invaluable help in the proofs of a theorem.