8-1-2004

Impartiality, Solidarity, and Priority in the Theory of Justice

Juan D. Moreno-Ternero

John E. Roemer

IMPARTIALITY, SOLIDARITY, AND PRIORITY IN THE THEORY OF JUSTICE

By

Juan Moreno-Ternero and John E. Roemer

August 2004

COWLES FOUNDATION DISCUSSION PAPER NO. 1477

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

http://cowles.econ.yale.edu/

Impartiality, solidarity, and priority in the theory of justice*

Juan D. Moreno-Ternero
Yale University

John E. Roemer
Yale University

July 30, 2004

Corresponding Author:
John E. Roemer.
Department of Political Science.
Yale University.
124 Prospect Street. PO Box 208301.
New Haven CT 06520-8301.
john.roemer@yale.edu

*We thank G.A. Cohen, Marc Fleurbaey, Shelly Kagan, Marco Mariotti, Klaus Nehring, Ariel Rubinstein, and Larry Temkin for helpful contributions to this paper. We also acknowledge the participants' contributions at seminars and conferences where the paper has been presented, at Rutgers University, the University of Ghent, the Ecole Polytechnique, Hitotsubashi University, and Osaka University. Moreno-Ternero's research was made possible thanks to a post-doctoral grant for advancement of research staff from the Generalitat Valenciana.

Abstract

The veil of ignorance has been used often as a tool for recommending what justice requires with respect to the distribution of wealth. We show that John Harsanyi's and Ronald Dworkin's conceptions of the veil, when modeled formally, recommend wealth allocations in conflict with the prominently espoused view that priority should be given to the worse off with respect to wealth allocation. It follows that those who believe that justice requires impartiality and priority must seek some method of assuring the former other than the veil of ignorance. We propose that impartiality and solidarity are fundamentals of justice, and study the relationship among these two axioms and priority. We characterize axiomatically resource allocation rules that jointly satisfy impartiality, solidarity, and priority: they comprise a class of general indices of wealth and welfare, including, as polar cases, the classical equal-wealth and equal-welfare rules.

JEL numbers: D63.

Keywords: Impartiality, Solidarity, Priority, Veil of ignorance, Allocation rules, Characterization result.

1 Introduction

The construct of the veil of ignorance (VI) has been of significant import in political philosophy during the last half century: three prominent writers–John Harsanyi, John Rawls, and Ronald Dworkin–have employed it in different forms. Although these three disagree on exactly how thick the veil should be, each uses it as a tool to enforce impartiality in the procedure that deduces what the worldly distribution of resources or wealth should be. The VI model is putatively impartial because the 'soul' or 'souls' or 'parties' or 'observer' who contemplate(s) behind the veil are (is) deprived of precisely that information that the author deems to be morally arbitrary, to use the phrase of Rawlsian parlance.

There are four important terms regarding veil-of-ignorance thought experiments, whose roles we wish to disentangle: impartiality, self-interest, rationality, and representation. For the veil-of-ignorance constructions that we study below, those of John Harsanyi and Ronald Dworkin, we use the following terminology: there are *persons* in the real world, represented by a *soul* or *souls* behind the veil. The persons are *self-interested*: each is concerned with only her own welfare, which is a function of only her own wealth. (Thus, the economic environment is classical.) A soul is *representative* and *rational*: representative in being a perfect agent of the person for whom it stands, and rational in the sense of recommending that wealth distribution which is in the best interest of its person, given the information it possesses. An allocation rule is *impartial* if it makes use only of information about persons that is morally relevant. The veil of ignorance is an *allocation rule*: it is a procedure that takes data about the real world as an input and produces a recommended allocation of resources as its output.

Thus self-interestedness is a property of worldly persons, rationality and representation (or, better perhaps, loyal agency) are properties of souls, and impartiality is a property of allocation rules.

We prefer this terminology to one saying that souls are impartial behind the veil of ignorance. A soul cares only about its person: in this sense it is not impartial–it is, indeed, very partial. Impartiality is better construed as a property of rules or institutions than of persons or souls.[1]

From quite a different vantage point, another group of political philosophers (which has a non-empty intersection with the first group) has been concerned to argue that justice requires that priority be given to the worse off. The most extreme form of priority is advocated by Rawls, for whom differences in amounts of primary goods accruing to people are only morally permissible if they maximize the level of (or index of) primary goods accruing to the worst off (that is, she who is least endowed with primary goods). Rawls (1971) attempts, unsuccessfully in our view, to argue for this principle using a veil-of-ignorance (original position) construction.[2]

The difference principle has often been criticized as being too extreme, and Derek Parfit (1997) has coined the term prioritarianism for the view that the worse off should be given priority over the better off with respect to resource allocation, but that they need not necessarily receive the extreme priority that characterizes maximin (the difference principle). In a welfarist setting, prioritarianism is usually characterized as a social welfare function with strictly convex upper contour sets. The boundaries of prioritarianism are maximin on one side, and utilitarianism on the other. (See, for example, Roemer (2004).)

Other philosophers who would identify themselves with either a prioritarian or egalitarian or difference-principle view include Brian Barry (1995), G.A. Cohen (1992), Thomas Scanlon (1998), and Larry Temkin (1993). There are surely many more. We include together the three views just mentioned because prioritarianism is a weakening of egalitarianism and the difference principle: if a rule is egalitarian or maximin it is surely prioritarian.[3]

[1] Note, in common parlance, we speak of 'impartial judges;' of course, the judge represents an institution.

[2] See Roemer (1996) for one discussion of the inadequacy of Rawls's argument from the original position. We also provide some discussion below in section 3.

[3] One could argue that egalitarianism does not imply priority, in the sense that $(2, 2)$ is more egalitarian than $(3, 4)$, but the worse off person is better off in the second allocation

Those who advocate priority but not maximin do so usually because they consider the costs of implementing the difference principle too great–costs borne by the better off.

In this paper, we show, first, that the veil of ignorance, when formulated in a rigorous way, is inconsistent with prioritarianism: to be precise, it will often recommend distributions of wealth that give priority to the better off.[4] If one insists that justice requires impartiality (or, more formally, that an allocation rule that implements justice must be impartial), which we and (probably) all others do, then it seems one must conclude from this demonstration that either justice is not prioritarian, or that the veil of ignorance improperly captures the kind of impartiality required of justice. We adhere to the view that justice is at least prioritarian, by which we mean not to exclude the difference principle, and in particular, the variant of it formulated in Cohen (1992).[5]

If the veil of ignorance is impartial in the sense that justice requires, and it is anti-prioritarian, must one conclude that justice is not prioritarian? One approach to escaping this inference would be to argue that the veil-of-ignorance does not model the *kind* of impartiality that justice requires. Pursuing this approach would require a careful conceptualization of impartiality, allowing one to delineate its several species. We do not take this approach here, though it perhaps could be fruitfully developed. At present we believe that, *qua* impartiality, the veil-of-ignorance thought experiment is just fine.

The alternative route, which we here pursue, is to argue that any al-

than in the first. Thus priority could recommend $(3, 4)$ but equality $(2, 2)$. One could, however, also argue that in $(2, 2)$ the first person is given greater priority than in $(3, 4)$. We pursue this no further.

[4] An early form of this work is available in Roemer (2002); that article has an error, which is corrected here.

[5] In Cohen's view, an allocation is not truly maximin if those who are better off could transfer wealth to those worse off, while still remaining better off than the latter after the transfer. That they selfishly may wish not to do so is not germane. Thus, (selfish) incentives do not justify income differences, as they do in Rawls's formulation.

location rule that delivers distributive justice must satisfy some important principle in addition to impartiality, and then to show that the *conjunction* of that new principle and impartiality excludes the veil of ignorance–more strongly, that their conjunction (intersection) implies prioritarianism. Indeed, our proposal is that the additional principle be one of *solidarity*.

We will delineate a formal model of wealth allocation, in an environment that poses the problem of distributive justice in a simple way. We will then propose what impartiality, solidarity, and priority require, and will show (at least for some domains) that the conjunction of impartiality and solidarity implies priority. Moreover, we will deduce precisely what wealth-allocation rules satisfy the axioms of impartiality, solidarity and priority.

A caveat is in order. We provide no argument that solidarity must be an axiom of justice, although there is a long history to the idea of solidarity, and perhaps to the view that justice requires it.[6] Our study of the veil of ignorance, and its anti-prioritarian consequences, leads us to suggest that *some other basic principle* besides impartiality is needed to characterize justice. We do not, to repeat, reject the view that the veil of ignorance is an impartial procedure for deciding upon the distribution of wealth: our method for excluding it, as the determinant of justice, is that it fails to satisfy *another* principle of justice.

Other principles that one might want to consider to append to impartiality, in lieu of solidarity, are fraternity or reciprocity (Liberté, fraternité, égalité). We leave such an investigation for another time.

Our stance places us in disagreement with the implication of the title of Brian Barry's book *Justice as Impartiality*. We do not believe that the kind of justice Barry wishes to derive can be shown to follow from impartiality alone, as his title suggests. We think that, if Barry deduces justice of the prioritarian variety, he must be smuggling in some assumption, like solidarity,

[6] We say 'perhaps,' because, in the labor movement, solidarity has certainly been a *strategy*, if not obviously an ethical canon. But the strategy has appeal, arguably, not only because it produces strength, but because it is morally *right*.

to do the work. We are probably at odds with the Kantian tradition as well, for arguably Kant believed that justice was characterized by impartiality and rationality. (The Kantian imperative is a statement of impartiality.) Similarly, if Scanlon (1998) deduces a prioritarian kind of justice, then the 'reasonableness' of proposals that plays the key role in his theory must, we claim, have embedded within it a conception of solidarity–or something like it to do the work that it does in our theory.

Our argument may be a disappointment for the advocates of what is called left-liberal political philosophy. For it has been a seductive goal of that school to show that prioritarian/egalitarian desiderata can be deduced from premises that will attract (almost) universal assent–premises such as rationality and impartiality. (There is a flavor of this in Rawls, although he is not totally clear on the issue.) We believe this goal is unachievable. To get equality/priority as a result, one must, we now believe, put something very close to it (like solidarity) in.

In the next section, we complete the theory of the veil of ignorance that Harsanyi began, and show its anti-prioritarian consequences. In Section 3, we model Ronald Dworkin's version of the veil of ignorance, and show its anti-prioritarian nature. In Section 4 we present an axiomatic theory of resource allocation involving the concepts of impartiality, solidarity, and priority. We characterize the allocation rules that satisfy these axioms as ones that equalize some index of wealth and welfare. Section 5 concludes. Finally, most of the proofs have been relegated to an Appendix.

2 The Harsanyi veil of ignorance

In 1953, John Harsanyi proposed the first precise model of the veil of ignorance. Suppose there are n individuals, each of whom possesses von Neumann-Morgenstern (vNM) preferences over wealth lotteries. Denote vNM utility functions on wealth for these people by $v^1, v^2, ..., v^n$. There is an amount of wealth $\overline{W}$ to be divided among them. What is the just division?

Harsanyi proposes to conceptualize a single impartial observer (IO) who will become one of these people, with equal probability of becoming each one. How would such an observer allocate the wealth?

The IO's data consist in the set $\{v^1, v^2, ..., v^n, \overline{W}\}$.

Denote by (i, W) the *extended prospect* that means 'becoming person i with wealth W.' Harsanyi proposes that the IO, to solve his problem, must itself possess a vNM utility function U defined on extended prospects. (That is, it must be able to evaluate lotteries on extended prospects.) We can then represent the 'birth lottery' through which the IO becomes a particular person, and in which the distribution of wealth among the individuals is $(W^1, W^2, ..., W^n)$, by

$$l = \left(\frac{1}{n} \circ (1, W^1), \frac{1}{n} \circ (2, W^2), ..., \frac{1}{n} \circ (n, W^n)\right).$$

This is to be read, "With probability $1/n$, the extended prospect $(1, W^1)$ is realized (and the IO becomes person 1 with wealth W^1), with probability $1/n$ the extended prospect $(2, W^2)$ is realized, and so on". Now the utility the IO receives from this lottery is, by the expected utility property, equal to:

$$\sum_{i=1}^{n} \frac{1}{n} \cdot U(i, W^i) \tag{1}$$

and so the IO need only find the distribution of wealth that maximizes expression (1) subject to the constraint that $\sum W^i = W$. That distribution is the one it would choose, and therefore, that justice recommends.

The problem, then, is to deduce what the function U is. Harsanyi takes an axiomatic approach to this problem. He assumes what he calls:

The Principle of Acceptance. *When contemplating wealth lotteries in which the individual i is fixed, the IO should accept the vNM preferences of individual i.*

Formally, this says:

For each fixed i, the function $U(i, \cdot)$ represents the same vNM preferences on wealth *lotteries as $v^i(\cdot)$ represents.*

Now the vNM theorem tells us that any two vNM utility functions that represent the same preferences must be positive affine transformations of each other. Therefore:

For all W and i, there exist $a^i > 0$ and b^i such that

$$U(i, W) = a^i \cdot v^i(W) + b^i \tag{2}$$

Substituting formulae (2) into (1), we have that

$$\sum_{i=1}^{n} \frac{1}{n} \cdot U(i, W^i) = \sum_{i=1}^{n} \frac{1}{n} \left(a^i \cdot v^i(W^i) + b^i \right) = \frac{1}{n} \cdot \sum_{i=1}^{n} a^i \cdot v^i(W^i) + \frac{1}{n} \cdot \sum_{i=1}^{n} b^i \tag{3}$$

Maximizing the right-hand side of (3) is equivalent to choosing that distribution of wealth that maximizes $\sum a^i \cdot v^i(W^i)$. That is the end of Harsanyi's argument: the IO must maximize *some* positive weighted sum of the vNM utilities of the individual persons.

But the argument is unfinished, for Harsanyi has provided no way of determining the values of the positive numbers $\{a^i : i = 1, 2, ...n\}$, so he has not determined the vNM preferences of the IO. Furthermore, there is *no way* to derive these values from the information that Harsanyi has provided to the IO.

A moment's thought will show why this is so. The only information the IO has, consists in the profile of risk preferences of the individuals, and the total wealth to be allocated. But to decide whether it would rather become Alan with \$1000 or Barbara with \$3000, the IO *must be able to compare how well off Alan is with* \$1000 *with how well off Barbara is with* \$3000. (Or it must have some independent reason to prefer to be Alan, say.) There is no way to avoid such comparisons, and there is no way the IO can make them with the information it has. There is, in Harsanyi's specification of the problem, *absolutely no information* permitting interpersonal welfare comparisons. The vNM preferences of the individuals are purely ordinal preferences that measure 'utility' in a non-comparable way across persons.[7]

[7] Many people are confused about this claim. VNM preferences are *ordinal* preferences

We propose what now appears to be the obvious move: to amend Harsanyi's model by providing the IO with additional information, allowing it to perform interpersonal welfare comparisons. To this end, we assume that there is a complete order on extended prospects, denoted $\succsim$. The statement $(1, W^1) \succsim (2, W^2)$ means 'person 1 with wealth W^1 is at least as well off as person 2 with wealth W^2'. The strict preference is denoted $\succ$ and welfare indifference is denoted $\sim$. This order is to be thought of as a fact about the world, a statement about how the persons experience life: it is not the subjective preference order of the IO.

We now append what we name:

The Principle of Neutrality. $U(i, W^i) \geq U(j, W^j) \Leftrightarrow (i, W^i) \succsim (j, W^j)$.

In other words, the IO weakly prefers one extended prospect to another if and only if the person in the first extended prospect experiences well-being at least as high as the person in the second extended prospect. We call this 'neutrality' because it asserts that the IO brings no external considerations to bear concerning what person it would like to become: it only follows the dictates of the interpersonally comparable attribute called welfare or well-being, ignoring all other traits these individuals have (such as their sex, race, nationality, religious preference, or political views).

Thus, the data available to the IO are now $\{v^1, v^2, ..., v^n, \overline{W}, \succsim\}$

We shall see that the two principles of Acceptance and Neutrality enable us completely to solve the problem of the IO's vNM preferences.

We first introduce another concept. Let $\{W_a^1, W_a^2, ..., W_a^n\}$ be an *equal-welfare distribution of wealth*: that is a distribution such that

$$(i, W_a^i) \sim (j, W_a^j) \text{ for every pair } i, j.$$

Let there be two more equal-welfare distributions of wealth denoted $\{W_b^i\}_{i=1}^n$ and $\{W_c^i\}_{i=1}^n$, and suppose that these three distributions of wealth represent

on lotteries. There happens to be a very useful *cardinal* representation of those preferences, which allows us to calculate the utility of a lottery in a very simple way (by factoring out the probabilities). But the *preferences* are purely ordinal and non-comparable across persons. For further discussion, see Roemer (1996, chapter 4).

three welfare levels in increasing order of welfare, and so it follows that for each i, $W_a^i < W_b^i < W_c^i$, because we assume that welfare is strictly increasing in wealth.

We again invoke the vNM theorem, which tells us that for each person i there is a unique probability p^i such that:

$$v^i(W_b^i) = p^i \cdot v^i(W_a^i) + (1 - p^i) \cdot v^i(W_c^i) \tag{4}$$

In general, of course, the probabilities p^i will differ across individuals. The more risk averse an individual is, the lower will p^i be. We say that the individuals in the world are *risk isomorphic* if, for any choice of the three equal-welfare distributions, the numbers $\{p^i : i = 1, \ldots, n\}$ are identical for all i. What this says is that, when viewing lotteries in terms of the *welfare* they provide to the individual in question, all individuals have identical risk preferences. Risk isomorphism is clearly a singular case, which will rarely if ever hold in 'real worlds.'

We have the following:

Theorem 1 *The following statements hold:*

(A) If the individuals in the world $\{v^1, v^2, ..., v^n, \overline{W}, \succsim\}$ are risk isomorphic, then there is a unique vNM preference order (for the IO) that satisfies the principles of acceptance and neutrality. This order is represented by the vNM utility function on extended prospects:

$$U(i, W) = \frac{v^i(W) - v^i(W_a^i)}{v^i(W_b^i) - v^i(W_a^i)}. \tag{5}$$

where $\{W_a^i\}_{i=1}^n$ and $\{W_b^i\}_{i=1}^n$ are any two equal-welfare distributions of wealth such that $W_b^i > W_a^i$.

(B) If the individuals are not risk isomorphic there is no vNM preference order on extended prospects that satisfies acceptance and neutrality.[8]

[8] This is a correction of the stated theorem in Roemer (2002). I (Roemer) there incorrectly assumed something that implied that all environments were risk isomorphic, and so it was claimed that the principles of neutrality and acceptance always characterized unique vNM preferences for the IO. Fortunately, the examples of that paper are correct, as they are all examples where risk isomorphism holds.

Proof.

We proceed in two steps to prove the theorem.

Step 1.

Suppose a vNM preference order on lotteries on extended prospects exists, satisfying Acceptance and Neutrality and let U be a vNM utility function representing it. Let $\{W_a^i\}_{i=1}^n$, $\{W_b^i\}_{i=1}^n$, $\{W_c^i\}_{i=1}^n$ be three equal-welfare wealth distributions such that $W_a^i < W_b^i < W_c^i$. Then, by Neutrality:

$$U(i, W_k^i) = U(j, W_k^j) \text{ for all } i, j = 1, ..., n \text{ and for all } k = a, b, c.$$

By Acceptance, there exist positive numbers α^i and numbers β^i and numbers K_a, K_b, K_c such that:

$$K_a = \alpha^i v^i(W_a^i) + \beta^i,\ K_b = \alpha^i v^i(W_b^i) + \beta^i \text{ and } K_c = \alpha^i v^i(W_c^i) + \beta^i \text{ for all } i.$$

We immediately have by subtracting these equations from each other:

$$\frac{K_b - K_a}{K_c - K_a} = \frac{v^i(W_b^i) - v^i(W_a^i)}{v^i(W_c^i) - v^i(W_a^i)} \text{ for all } i. \tag{6}$$

Let the fractions p^i be defined uniquely by:

$$v^i(W_b^i) = p^i \cdot v^i(W_c^i) + (1 - p^i) \cdot v^i(W_a^i)$$

Rearrange to show that:

$$p^i = \frac{v^i(W_b^i) - v^i(W_a^i)}{v^i(W_c^i) - v^i(W_a^i)}. \tag{7}$$

It follows from (6) and (7) that for all i, we must have $p^i = \frac{K_b - K_a}{K_c - K_a}$, a constant. Hence risk isomorphism is necessary for the existence of a vNM preference order satisfying Acceptance and Neutrality. In particular, this proves part (B) of the theorem.

Step 2.

Conversely, suppose risk-isomorphism holds. Define U as in (5). Clearly, Acceptance holds: for each i, $U(i, \cdot)$ is a positive affine transformation of v^i. We must also show that U represents the interpersonal ordering.

Let $(i, W) \sim (j, W')$. We compute:

$$U(i, W) = \frac{v^i(W) - v^i(W_a^i)}{v^i(W_b^i) - v^i(W_a^i)}, \text{ and } U(j, W') = \frac{v^j(W') - v^j(W_a^j)}{v^j(W_b^j) - v^j(W_a^j)}.$$

Risk isomorphism implies that these two values are equal, so $U(i, W) = U(j, W')$. Suppose, now, that $(i, W) \succ (j, W')$. Define W^* by $(i, W^*) \sim (j, W')$. We now know that $U(i, W^*) = U(j, W')$. But since $v^i(W) > v^i(W^*)$, substitution into the definition of $U(i, W)$ immediately shows that $U(i, W) > U(j, W')$. This demonstrates the claim.

If we take two other equal-welfare wealth distributions from the ones chosen here, call them $\{\widehat{W}_a^i\}$ and $\{\widehat{W}_b^i\}$, it is a simple algebraic exercise to show, by invoking risk isomorphism, that the new function, call it $\widehat{U}$, thereby defined, is an affine transformation of the function U. Thus, the vNM preferences of the IO are well-defined, independent of the choice of equal-welfare wealth distributions. ■

Part B of the theorem is an impossibility theorem. It says that, in what is the usual case (of risk non-isomorphism), the Harsanyi veil of ignorance, amended by the principle of neutrality, is an incoherent thought experiment. In the singular case of risk-isomorphism, we uniquely determine the preferences of the IO (that is, we solve for the coefficients $\{a^i\}$ of equation (3).)

Let us examine the implications of part A with a simple example. There are two individuals, Andrea and Bob. They are each risk neutral. We may therefore take them to have the same linear vNM utility function, namely

$$v^A(W) = v^B(W) = W.$$

Let us suppose that the interpersonal welfare order is given by $(Andrea, W) \sim (Bob, 2W)$; that is, Bob always needs twice the wealth of Andrea to achieve the same welfare level as she. It is easy to see that this environment is risk isomorphic.

We now compute what the IO recommends under the preferences of part A of the theorem. Suppose that $\overline{W} = 1$, so a distribution of wealth is represented by $(W, 1-W)$ where the first component goes to Andrea and the second to Bob. The IO must choose W. We know that $U(A, W) = U(B, 2W)$ by the principle of neutrality. Now the IO must choose W to

$$\text{maximize } \frac{1}{2}U(A, W) + \frac{1}{2}U(B, 1 - W).$$

By the formula just given we can write this as

$$\max \frac{1}{2}U(A, W) + \frac{1}{2}U(A, \frac{1 - W}{2})$$

But by the principle of acceptance, this is equivalent to maximizing

$$\frac{1}{2}W + \frac{1}{2} \cdot \frac{1 - W}{2} = \frac{1}{4} + \frac{W}{4} \tag{8}$$

which is achieved at $W = 1$: the IO would give all the wealth to Andrea.

Now in this environment, we consider Bob to be *disabled* with respect to Andrea: he requires more wealth than she to receive any given level of welfare. Thus, the Harsanyi VI gives all the wealth to the able person.

The general result is: *If all individuals are risk neutral, and they can be ordered with respect to 'ability', the talent of converting wealth into welfare, then the Harsanyi VI assigns all the wealth to the most able individual(s).*

Of course, the interpretation matters here. A situation where Bob requires twice Andrea's wealth to reach her level of welfare could also be due to Bob's having expensive tastes for which we hold him responsible, and in that case, we might not be so disturbed by the conclusion.[9] But we insist that that is not the problem we are here studying. We are discussing worlds where people differ in their ability to convert wealth into well-being, through no fault of their own.

We say that *priority requires that disabled individuals receive at least as much wealth as able ones.* That is our definition of priority for these worlds. In the environments under discussion, we have a clear way of deciding what 'being worse off' means: it means 'requiring more wealth than another to reach any given welfare level'.

[9]The issue of expensive tastes is focal in the contemporary literature on distributive justice: see, for the locus classicus, Dworkin (1981a).

What happens if we alter the risk preferences in the above example so that the individuals are risk averse? For small degrees of risk aversion, it continues to be the case that the amended Harsanyi veil delivers more wealth to the able agent, although it will deliver some wealth to both agents. Only for large degrees of risk aversion does the veil of ignorance assign more wealth to the disabled person.

Now consider part B of the theorem. We have no vNM preferences for the IO in the case of risk non-isomorphism. However, we propose the following procedure. Denote the individuals by $1, 2, \ldots, n$. The IO first takes on the vNM preferences of any person i, and chooses the wealth distribution i would choose, if she always converts wealth given to other people into the welfare-equivalent wealth for herself (i). We define this precisely as follows. For any pair (j, W^j) and any agent i define W_i^j by $(j, W^j) \sim (i, W_i^j)$. That is, W_i^j is the wealth that i would have to have to reach the same level of welfare as j achieves with wealth W^j. We assume that it is always possible to find such a wealth level–that is, welfares can always be equalized across persons, were sufficient wealth available.[10] If the distribution of wealth being contemplated is $(W^1, W^2, ..., W^n)$ then the IO, placing herself in i's shoes, would evaluate the birth lottery as having expected utility

$$\sum_{j=1}^{n} \frac{1}{n} \cdot v^i(W_i^j). \tag{9}$$

Thus the IO, using i's risk preferences, asks how she would feel as any person j, given the wealth j gets in the distribution: to do so, the IO must convert j's wealth to the welfare–equivalent wealth for i, since the IO is evaluating everything from i's perspective.

Harsanyi used the phrase *extended sympathy* for the compassion the IO feels as it contemplates being different people. But, since Harsanyi did not deal with interpersonal comparisons of welfare, his extended sympathy only referred to the IO's taking on the risk preferences of different people, as

[10] We exclude from consideration hard cases, where the upper bounds of welfare achievable by persons differ.

modeled by his Principle of Acceptance. The formulation (9) is one of *truly extended sympathy.* The IO, when stepping in the shoes of person i, imagines how i would feel if he were to be realized as any person j with a given wealth level W^j. He would experience j's wealth level W^j, which is equivalent to i's having the wealth level W_i^j.

Denote by $\omega^i = (\omega_1^i, ..., \omega_n^i)$ the feasible distribution of wealth that maximizes expression (9) subject to the condition that $\sum_{k=1}^n \omega_k^i = W$. Sequentially, the IO now performs this computation, taking on every person's viewpoint. This produces n wealth distributions $\omega^1, ..., \omega^n$. We propose that the IO take the average of these distributions, $\frac{1}{n} \cdot \sum \omega^i$, as her recommended distribution.[11]

This procedure can be performed for any environment, risk isomorphic or not. It is, furthermore, a *generalization* of the procedure of part (A) of Theorem 1: that is, it is easy to demonstrate that if the environment is risk-isomorphic, the procedure just described coincides in its recommendation with maximizing the IO's vNM utility function. This follows from the fact that in the case of risk-isomorphism, the n wealth distributions ω^i are all identical, and each is the distribution recommended by maximizing the utility function in part (A) of the theorem.

It therefore follows that our general procedure is also anti-prioritarian, because in the special case of risk-isomorphism, we know it is anti-prioritarian.

We have now provided the argument that the veil of ignorance, properly completed from Harsanyi's important first step, is anti-prioritarian, in the sense that it fails in general to assign at least as much wealth to disabled agents as to able ones. Although Harsanyi's assumption that the IO must possess vNM preferences is too strong–in the sense that no such preferences exist that satisfy the very reasonable axioms of acceptance and neutrality except in a singular case–we have produced an attractive proposal for what the IO should do in the general case, and it also is anti-prioritarian.[12]

[11] Indeed, as we will see shortly, the IO can take any convex combination of these wealth distributions.

[12] One might try to defend Harsanyi's veil of ignorance *and* prioritarianism by saying

3 The Rawls and Dworkin veils of ignorance

Rawls (1971) deprives the souls (whom he calls 'parties') in the original position of knowledge of the preferences of those persons they represent, and even of how preferences are distributed. We believe there is no way of solving the problem for the Rawlsian soul (for, as many have remarked, there really is only one soul). That soul faces virtually complete ignorance. Certainly Rawls provides no coherent argument.

We do not accept Rawls's argument for why souls behind the veil should not know the distribution of preferences. Rawls wishes the parties to think only in terms of primary goods: but if Rawls wants justice to concern itself with primary goods, and if he wants a veil of ignorance to produce justice, then a concern with primary goods is what souls behind his veil should *end up* with, when they are concerned with their own welfare maximization. Presumably, the argument should be that bargaining behind the veil, among persons with disparate preferences, will produce an agreement to focus upon primary goods. Rawls stacks the deck for his view, however, by depriving souls of knowledge of preferences.

If preferences are morally arbitrary, then souls should at least know the *distribution* of preferences. This Rawls denies them because he fears that a soul would be partial to a preference order that has a high frequency. But this, we think, is an incorrect invocation of impartiality. First, we have pointed out that there is no inconsistency in souls being partial and the *veil of ignorance* being an impartial allocation rule. Second, there is an automatic mechanism that prevents a decision maker from giving too much wealth to those with preference orders (or utility functions) that appear with

that, when such monumental issues are at stake as one's wealth for a lifetime, rational individuals would be highly risk averse, thus excluding from the domain of possible worlds, profiles of risk preferences which generate the conflict with priority. We are skeptical. Real people frequently take life-threatening risks that indicate that they do not have excessively high degrees of risk aversion. It is unappealing to say that the only rational persons are the ones who are extremely risk averse.

a high frequency: horizontal equity requires she gives the same amount to all people with that utility function, and since there are many such people, by hypothesis, she cannot give too much wealth to each of them. This suffices to control the decision maker who might be 'partial' to those with common preferences.

On the other hand, if preferences are not morally arbitrary (if, for instance, persons are to held responsible for their plans of life, as Rawls sometimes says), then the souls behind the veil should know those preferences. The original position is only meant to shield souls from knowledge of attributes that are morally arbitrary.

Ronald Dworkin (1981b), in contrast to Rawls, has outlined a conception of the veil of ignorance that is coherent and can be modeled formally. Here, we present a simple two-person version, which suffices for our purposes. (For a more leisurely discussion of Dworkin's insurance mechanism behind the veil of ignorance, see Roemer (1996, chapter 7).)

Suppose we again have Andrea and Bob, and Bob is disabled with respect to Andrea–to wit, he requires $2W$ in wealth to reach the same welfare level as Andrea reaches with W. For the sake of variety, we will now suppose that Andrea and Bob have the same risk preferences over wealth and their vNM utility function is given by

$$v(W) = \sqrt{W}.$$

This time, Andrea and Bob are risk averse.

Dworkin wishes to hold persons responsible for their risk preferences, but not for their talents. In this case, the talent is the ability to convert wealth into welfare. Thus, behind the veil of ignorance he constructs, the soul representing a person knows its person's vNM utility function, but does not know its person's talent.

Behind the veil of ignorance, there are two souls–call them α and β–who represent Andrea and Bob, respectively. Each soul knows the welfare producing capacities of Andrea and Bob, and each believes that it will become Andrea or Bob with equal probability (or, to paraphrase, that it will acquire

Andrea's and Bob's talent with equal probability).

Thus there are two states of the world, from the viewpoint behind the veil, as follows:

State	α becomes	β becomes
1	Andrea	Bob
2	Bob	Andrea

In state 1, soul α becomes Andrea and soul β becomes Bob; in state 2, the assignments of souls to persons are the other way around. We know that state 1 will occur, but the souls behind the veil assign a probability of one-half to each state's occurring.

We assume that, in the real world, Andrea has an endowment W^A of wealth and Bob has an endowment of W^B.

Behind the veil, the souls purchase insurance against bad luck in the birth lottery. We assume (after Dworkin) that the souls have equal purchasing power for insurance. This is where equality enters importantly into Dworkin's view. It doesn't matter how much purchasing power they each have: we shall say each has zero. This means that the only way to purchase insurance for indemnity in one state is to sell insurance for the other's indemnity in the other state.

We model the insurance market as follows. There are two commodities: the first is a contract which will deliver \$1 to the holder should state 1 occur, and the second is a contract which will deliver \$1 to the holder should state 2 occur. Let us denote the prices (behind the veil) for these two commodities by p^1 and p^2. Note that these commodities are purchased behind the veil, using the currency that exists there (clamshells, to follow Dworkin), not worldly wealth.

Denote by x_1^α and x_2^α the amount of commodity 1 and commodity 2, respectively, that soul α purchases. If x is positive, that means she purchases contracts that will deliver to her x dollars if the state of the subscript occurs; if x is negative, that means she will deliver x dollars to someone else should that state occur. The budget constraint for soul α is

$$p^1 x_1^\alpha + p^2 x_2^\alpha = 0,$$

which means that the amount of commodity 1 she can purchase must cost exactly the income she generates by selling commodity 2 (or, the other way around). This constraint derives from the fact that her endowment of 'clamshells' behind the veil is zero. If the soul faces prices (p^1, p^2) then her optimization problem is as follows: choose x_1^a and x_2^a to maximize

$$\frac{1}{2}\sqrt{W^A + x_1^\alpha} + \frac{1}{2}\sqrt{\frac{W^B + x_2^\alpha}{2}} \text{ subject to } p^1 x_1^\alpha + p^2 x_2^\alpha = 0. \tag{10}$$

The objective she maximizes is her expected utility. The expression under the first radical is clear: this is what her wealth will be if she becomes Andrea (state 1). The expression under the second radical is trickier. In state 2, she becomes Bob; the wealth she (he) would then have is $W^B + x_2^\alpha$. However, she must evaluate this wealth from Andrea's viewpoint–and by hypothesis the welfare this amount of wealth generates for Bob is exactly the welfare that one-half this amount generates for Andrea. So truly extended sympathy gives us the objective in (10).

In like manner, the optimization problem for soul β is to choose x_1^β and x_2^β to maximize

$$\frac{1}{2}\sqrt{W^B + x_1^\beta} + \frac{1}{2}\sqrt{2(W^A + x_2^\beta)} \text{ subject to } p^1 x_1^\beta + p^2 x_2^\beta = 0.$$

Note that, if soul β becomes Andrea, she must evaluate her wealth in terms of the welfare-equivalent wealth for Bob.

An *equilibrium* in the insurance market consists in:

(1) a pair of prices p^1 and p^2, and

(2) commodity demands $(x_1^\alpha, x_2^\alpha, x_1^\beta, x_1^\beta)$ such that the markets for both commodities clear, that is: $x_1^\alpha + x_1^\beta = 0 = x_2^\alpha + x_2^\beta$.

There is a unique equilibrium[13] in this market. It is:

$$\begin{aligned}
p^1 &= p^2 = 1 \\
x_1^\alpha &= \frac{2W^B - W^A}{3},\ x_2^\alpha = -x_1^\alpha \\
x_1^\beta &= \frac{W^A - 2W^B}{3},\ x_2^\beta = -x_1^\beta.
\end{aligned}$$

As we said, we know that, in the event, state 1 occurs; this means that the final wealth levels (under the Dworkinian tax scheme) must be

$$\begin{aligned}
W^{A,final} &= W^A + x_1^\alpha = \frac{2}{3} \cdot (W^B + W^A) \\
W^{B,final} &= W^B + x_1^\beta = \frac{1}{3} \cdot (W^B + W^A)
\end{aligned}$$

Thus, disabled Bob ends up with one-third of the total wealth, and able Andrea ends up with two-thirds of the total wealth.

In other words, the Dworkinian insurance market is in *general* anti-prioritarian. It does *not* (in general) assign at least as much wealth to the disabled person as to the able person.

This section and the last one do not prove that veils of ignorance are necessarily anti-prioritarian: we have established, however, that the two most coherent proposals in the philosophical literature for conceptualizing the veil of ignorance are so.[14]

4 Axiomatics

In this section, we introduce a formal framework in which the various concepts that have appeared above are central–interpersonal comparability of welfare, ability and disability, priority, and impartiality. We also introduce formally solidarity as a property of allocation rules. We have two purposes: (1) to see whether priority can be deduced from impartiality and solidarity,

[13] To be precise, the demands and supplies are uniquely determined. The prices can be any pair of equal positive numbers.

[14] We have not studied Binmore's (1994) formulation of the veil.

as we discussed in Section 1, and (2) to characterize those allocation rules which jointly satisfy impartiality, solidarity, and priority. It will turn out that our characterization echoes a theme that has appeared occasionally in the philosophical literature.

In the present formalization, persons' risk preferences do not appear in the description of possible worlds. We model a domain of environments where a social endowment of a resource must be allocated among individuals who have different capacities for transforming wealth into interpersonally comparable welfare. We will comment below on the absence of risk preferences.

4.1 The Model

Let $\mathbb{I}$ represent a population of all potential individuals (a set with an infinite number of members) and let $\mathcal{I}$ be the family of all finite subsets of $\mathbb{I}$. An element $I \in \mathcal{I}$ describes a finite set of individuals. Individuals derive welfare from a resource, called *wealth*. We assume that $\mathbb{I} \times \mathbb{R}_+$ is endowed with a complete order. As in Section 2, the expression $(i, W) \succsim (j, W')$ is read: "individual i equipped with wealth W enjoys a welfare level at least as high as individual j equipped with wealth W'". We assume that this order is continuous in W, and satisfies that, for any $i, j \in I$ and $W \in \mathbb{R}_+$ there is a wealth level such that $(i, W) \sim (j, W')$. We further assume that for any pair $i, j \in I$, $(i, 0) \sim (j, 0)$. A wealth level of zero can be thought of as inducing death, which is an equally bad outcome for all individuals. We assume that welfare is strictly increasing in wealth for every individual. Finally, we assume that, for any individual, very high welfare levels can only be achieved with very high wealth levels.

It is convenient to represent this interpersonally level comparable welfare ordering as follows. Fix a particular individual and call her individual 0. For any other individual i define a function $\sigma_i : \mathbb{R}_+ \to \mathbb{R}_+$ where for each $W \in \mathbb{R}_+$, $\sigma_i(W)$ is such that

$$(0, W) \sim (i, \sigma_i(W))$$

In other words, $\sigma_i(W)$ is the wealth that i must receive in order that she enjoys the same level of welfare as individual 0 enjoys with wealth W.[15] The assumptions on $\succsim$ tell us that for all i, σ_i is a continuous strictly increasing unbounded function such that $\sigma_i(0) = 0$.

We say that a family of functions constitutes a dense domain if the graphs of these functions cover the positive quadrant. We shall also assume that $\{\sigma_i : i \in \mathbb{I}\}$ constitutes a dense domain. Formally,

Dense Domain *(DD). $\{\sigma_i : i \in \mathbb{I}\}$ is a dense domain, i.e., for every $(a, b) \in \mathbb{R}^2_{++}$ there exists an individual $i \in \mathbb{I}$ such that $\sigma_i(a) = b$.*

We now introduce the concept of ability in relative terms. We say that an individual is more able than another one if the former needs less wealth than the latter one to reach the same level of welfare. Formally,

*An individual i is **able** with respect to an individual j if $\sigma_i \leq \sigma_j$ and $\sigma_i \neq \sigma_j$. We also say that, in this case, individual j is **disabled** with respect to individual i.*

Two individuals are **comparable** if one is at least as able as the other. Obviously, there might be individuals that are not comparable.

Note that the functions $u_i = (\sigma_i)^{-1}$ comprise a profile of utility functions for individuals which measure utility in a level comparable way. First, the functions u_i are defined on $\mathbb{R}_+$, since the σ_i functions are unbounded and strictly increasing. Second, they satisfy the following.

Claim. $(i, W_i) \succsim (j, W_j) \Leftrightarrow u_i(W_i) \geq u_j(W_j)$.
Proof of the claim. Let $k_i = u_i(W_i)$ and $u_j(W_j) = k_j$. Then individual i with wealth W_i enjoys the welfare that individual 0 receives with wealth k_i and individual j with wealth W_j enjoys the welfare that individual 0 enjoys with wealth k_j. Then, $(i, W_i) \sim (0, k_i)$ and $(j, W_j) \sim (0, k_j)$. Since welfare is strictly increasing in wealth for every individual, it follows that $(0, k_i) \succsim (0, k_j)$ if and only if $k_i \geq k_j$. This demonstrates the claim. ■

[15] In particular, σ_0 is the indentity function.

In particular, an individual i is able with respect to an individual j if and only if $u_i \geq u_j$ and $u_i \neq u_j$. It is also worth noting that the properties of the σ_i functions ensure that the u_i functions are continuous, strictly increasing, satisfy $u_i(0) = 0$ and constitute a dense domain. In particular, they are unbounded, i.e., $\lim_{x\to\infty} u(x) = \infty$. For the sake of completeness, we denote by $\mathcal{U}$ the set of functions satisfying these properties, i.e.,

$$\mathcal{U} = \{u : \mathbb{R}_+ \to \mathbb{R}_+ : \text{continuous, strictly increasing and s.t. } u(0) = 0,\ \lim_{x\to\infty} u(x) = \infty\}$$

We define an **economy** e as a triple (I, u, W), where $I \in \mathcal{I}$ is the set of individuals, $u = (u_i)_{i\in I} \in \mathcal{U}^{|I|}$ is the profile of utility functions (defined as above) for individuals in I, and $W \in \mathbb{R}_+$ represents the available wealth. The family of all economies is $\mathcal{E}$.

Consider a dense (in the sense of DD) sub-domain of agents $\{\sigma_i : i \in \mathbb{I}^c\}$ such that any two agents within it are comparable with respect to ability. Clearly there are many such sub-domains $\mathbb{I}^c$ of $\mathbb{I}$. Denote by $\mathcal{E}^c$ the set of all economies with wealth non-negative real and whose members comprise a finite subset of $\mathbb{I}^c$, i.e.,

$$\mathcal{E}^c = \{(I, u, W) \in \mathcal{E} : I \subset \mathbb{I}^c \text{ and } I \text{ finite}\}.$$

An **allocation rule** is a function F that associates to each economy $e = (I, u, W) \in \mathcal{E}$ a unique point $F(e) = (F_i(e))_{i\in I} \in \mathbb{R}_+^{|I|}$ such that $\sum_{i\in I} F_i(e) = W$. That is, an allocation rule indicates how to distribute the wealth available in an economy among its members. We shall present several characterization results for allocation rules. Some of the results will be proven on the whole domain of economies $\mathcal{E}$ whereas others will be proven for the restricted domain $\mathcal{E}^c$. We begin by introducing three axioms that model *impartiality*. First, two domain axioms:

Universal Domain *(UD). The allocation rule F is defined on the class of economies $\mathcal{E}$.*

Restricted Domain *(RD). The allocation rule F is defined on the class of economies $\mathcal{E}^c$.*

The domain axioms model impartiality because they exclude much information about persons that we consider ethically irrelevant. We next introduce a weak form of anonymity, namely:

Horizontal Equity *(HE). Let* $e = (I, u, W) \in \mathcal{E}$ *and* $i, j \in I$ *such that* $u_i = u_j$*. Then* $F_i(e) = F_j(e)$*.*

HE is clearly an axiom of impartiality.

The next axiom, consistency, is not motivated by one of our basic principles. It says that if a sub-group of two individuals secedes with the resource allocated to it under F, then in the smaller economy F allocates the resource in the same way.[16] The reader is referred to Young (1994), Roemer (1996) or Thomson (2004) for the many applications that exist in the literature on distributive justice concerning this notion.

Consistency *(CY). Let* $e = (I, u, W) \in \mathcal{E}$*. Let* $I' \subset I$ *with* $|I'| = 2$ *and* $e' = (I', u', W')$*, where* $u' = (u_i)_{i \in I'}$ *and* $W' = \sum_{i \in I'} F_i(e)$*. Then* $F_i(e) = F_i(e')$*, for all* $i \in I'$*.*

Because we provide no ethical basis for CY, we should eventually eliminate it from our evaluation. But it is useful to formulate it for the interim.[17]

We turn now to *solidarity*. Here we rely upon a literature which has formulated various solidarity axioms in the past twenty years.[18] Our foun-

[16] This axiom is usually referred in the literature as *bilateral consistency* to distinguish it from the stronger axiom in which no restriction on the cardinality of the seceding group is imposed. Since our study only requires the weaker axiom, and for ease of notation, we avoid the word bilateral.

[17] Thomson and Lensberg (1989) motivate consistency by saying that for a distribution to be just in society, it must be just in every sub-society. We do not wish to rely on this view.

[18] Alternative versions of solidarity have been considered in different contexts like fair division (e.g., Thomson, 1983; Roemer, 1986), surplus-sharing (e.g., Keiding and Moulin, 1991) collective choice (e.g., Sprumont, 1996; Ehlers and Klaus, 2001), social choice (e.g., Chun, 1986), house allocation (e.g., Ehlers and Klaus, 2004), assignment (e.g., Klaus and Miyagawa, 2001) or production models in which agents are assumed to have unequal production skills (e.g., Fleurbaey and Maniquet, 1999). For further details, the reader is

dational concept of solidarity says that, when a bad or good shock comes to an economy, all its members should share in the calamity or windfall. The first example of this ethos was 'resource monotonicity', introduced in Roemer (1986). Here, we write it as:

Resource monotonicity *(RM). Let* $e = (I, u, W)$ *and* $e' = (I, u, W') \in \mathcal{E}$ *such that* $W > W'$. *Then* $F_i(e) > F_i(e')$ *for all* $i \in I$.

Another solidarity axiom that we consider is that of *agent monotonicity.* It can be interpreted as saying that, if an individual 'becomes ill', either he receives no more resource than before, or he receives more resource, and in the latter case, all other individuals chip in to fund the additional amount.

Agent monotonicity *(AM). Let* $e = (I, u, W)$ *and* $e' = (I', u', W) \in \mathcal{E}$, *such that* $I = I_0 \cup \{j\}$, $I' = I_0 \cup \{j'\}$ *and* $u_j \geq u_{j'}$. *Then either* $F(e') = F(e)$, *or* $F_{j'}(e') > F_j(e)$ *and* $F_i(e') < F_i(e)$ *for all* $i \in I_0$.

The third solidarity axiom that we introduce is that of *agreement.* It says that the arrival of immigrants, whether or not accompanied by changes in the available wealth, should affect all original agents in the same direction: all gain or all lose, or all receive the same as before. Two similar axioms were considered by Chun (1999) under the name of "agreement" and "population and resource monotonicity", in the context of bankruptcy problems.

Agreement *(AG). Let* $e = (I, u, W)$ *and* $e' = (I', u', W') \in \mathcal{E}$, *such that* $I' \subseteq I$. *Let* $F_{I'}(e)$ *denote the projection of* $F(e)$ *onto the set of coordinates corresponding to* I'. *Then either* $F(e') = F_{I'}(e)$, $F(e') > F_{I'}(e)$ *or* $F(e') < F_{I'}(e)$.

It is straightforward to show that AG implies RM. It also implies CY. Conversely, if CY is extended to any subgroup of agents (not necessarily with cardinality 2) then CY and RM together imply AG (see the Appendix for the details).

We next propose axioms of *priority.* Following the discussion of Section

referred to Thomson (1995) and the literature cited therein.

1, we say an allocation rule is semi-prioritarian if it awards disabled agents at least as much wealth as able agents.

Semi-Priority *(SP). Let* $e = (I, u, W) \in \mathcal{E}$ *and* $i, j \in I$ *such that* $u_i \geq u_j$*. Then* $F_i(e) \leq F_j(e)$*.*

This is referred as the *weak equity axiom* by Sen (1973).

Semi-Priority guarantees that disabled agents receive at least as much wealth as abler ones: we discriminate positively towards the disabled. The motivation for the prioritarian view, however, also contains a negative element: society does not necessarily owe disabled agents as much as they would receive in a maximin allocation -that is, its obligation towards the unfortunate is limited. We attempt to capture this view with the following axiom.

Limited priority *(LP) Let* $e = (I, u, W) \in \mathcal{E}$ *and* $i, j \in I$ *such that* $u_i \geq u_j$*. Then* $u_i(F_i(e)) \geq u_j(F_j(e))$*.*

In other words, a disabled person should never be resourced to the extent that her welfare exceeds that of an able agent.

We define **prioritarianism** as the conjunction of semi and limited priority.

We conclude with a stronger notion of prioritarianism which makes no mention of relative ability. It says that no agent can dominate another agent both in resources and welfare.

Strong Priority *(TP) Let* $e = (I, u, W) \in \mathcal{E}$ *and* $i, j \in I$ *such that* $F_i(e) < F_j(e)$*. Then* $u_i(F_i(e)) \geq u_j(F_j(e))$*.*

As shown in the next result, TP implies both SP and LP. In $\mathcal{E}^c$, SP and LP together turn out to be equivalent to TP.

Proposition 1 *The following statements hold:*

(i) TP *implies* SP *and* LP*.*

(ii) In $\mathcal{E}^c$*,* SP *and* LP *together imply* TP*.*

Proof.

(i) Let F be an allocation rule satisfying TP.

If, contrary to the claim, F does not satisfy SP, then there exists $e = (I, u, W) \in \mathcal{E}$ and $i, j \in I$ such that $u_i \geq u_j$ and $F_i(e) > F_j(e)$. In particular, and since u_i is increasing, $u_i(F_i(e)) \geq u_j(F_i(e)) > u_j(F_i(e))$, which contradicts TP.

If, contrary to the claim, F does not satisfy LP, then there exists $e = (I, u, W) \in \mathcal{E}$ and $i, j \in I$ such that $u_i \geq u_j$ and $u_i(F_i(e)) < u_j(F_j(e))$. By TP, it follows that $F_i(e) \geq F_j(e)$. By SP, it follows that $F_i(e) \leq F_j(e)$. Thus, $F_i(e) = F_j(e)$ and therefore $u_i(F_i(e)) \geq u_j(F_j(e))$, which is a contradiction.

(ii) Suppose now that F is an allocation rule satisfying SP and LP.

Let $e = (I, u, W) \in \mathcal{E}^c$ and $i, j \in I$ such that $F_i(e) < F_j(e)$. Then, by SP, $u_i \geq u_j$. Thus, by LP, $u_i(F_i(e)) \geq u_j(F_j(e))$, as desired. ■

It is straightforward to show that SP implies HE. So does LP. Thus, in virtue of Proposition 1 it also follows that TP implies HE. Conversely, as shown in the next result, if a rule satisfies HE and AM then it also satisfies SP.

Proposition 2 *HE and AM together imply SP.*

Proof.

Let F be an allocation rule satisfying HE and AM. Suppose, contrary to the claim, that F does not satisfy SP. Then, there exists an economy $e = (I, u, W) \in \mathcal{E}$ and two agents i and $j \in I$ such that $u_i \geq u_j$ and $F_i(e) > F_j(e)$. Consider the economy e' in which agent j is replaced by an agent l identical to i. By AM, $F_l(e') \leq F_j(e)$ and $F_k(e') \geq F_k(e)$ for all $k \in I \setminus \{j\}$. By HE, $F_i(e') = F_l(e')$. Therefore $F_i(e) \leq F_i(e') = F_l(e') \leq F_j(e)$, a contradiction. ■

Proposition 2 is a simple demonstration that impartiality and solidarity imply (semi)priority. We are, however, interested in a sharper characterization.

4.2 A family of rules

We now construct a family of allocation rules. To do so, let Φ be the class of functions composed of all functions $\varphi : \mathbb{R}^2_{++} \cup \{(0,0)\} \to \mathbb{R}_+$, continuous on its domain and non-decreasing, such that $\inf\{\varphi(x,y)\} = \varphi(0,0) = 0$ and for all $(x,y) > (z,t)$, $\varphi(x,y) > \varphi(z,t)$.

Let φ be a function in the class Φ.[19] Let $S(e)$ denote the simplex of wealth allocations which are feasible in the economy $e = (I,u,W) \in \mathcal{E}$, i.e., $S(e) = \{x = (x_i)_{i \in I} \in \mathbb{R}^{|I|}_+ : \sum x_i = W\}$. Define the allocation rule, called the L_φ rule, by:

$$L_\varphi(e) = \mathit{leximin}\{\varphi(x_i, u_i(x_i)) : x = (x_i)_{i \in I} \in S(e)\}.$$

This means that $L_\varphi(e)$ is the wealth allocation that lexicographically maximizes the φ value across individuals in e. Note that applied in this manner to an agent's wealth and welfare, φ can be considered to be a generalized index of wealth and welfare. So the rules just defined leximin a generalized index of wealth and welfare.[20]

We have the following:

Proposition 3 *Let $\varphi \in \Phi$. For each $e = (I,u,W) \in \mathcal{E}$,*

$$L_\varphi(e) = \mathit{maximin}\{\varphi(x_i, u_i(x_i)) : x = (x_i)_{i \in I} \in S(e)\}.$$

Proof.

Let $e = (I,u,W) \in \mathcal{E}$ be given and denote $x = (x_i)_{i \in I} = L_\varphi(e)$. We show that $\varphi(x_i, u_i(x_i)) = \varphi(x_j, u_j(x_j))$ for all $i, j \in I$. Suppose, by contradiction, that $\varphi(x_i, u_i(x_i)) > \varphi(x_j, u_j(x_j))$. Since $\inf\{\varphi(x,y)\} = \varphi(0,0) = 0$, then $x_i > 0$. Thus, by continuity of φ, u_i and u_j, it follows that

$$\varphi(x_i - \varepsilon, u_i(x_i - \varepsilon)) > \varphi(x_j + \varepsilon, u_i(x_j + \varepsilon)),$$

[19] It should be noted that a function $\varphi \in \Phi$ cannot in general be extended to a continuous function on $\mathbb{R}^2_+$.

[20] We are indebted in a major way to Klaus Nehring, who suggested the L_φ rules.

for some $\varepsilon > 0$ sufficiently small. This contradicts the premise that we have leximinned φ. ■

The family $\{L_\varphi\}_{\varphi\in\Phi}$ is characterized by the axioms introduced in Section 4.1, as the next result shows.

Theorem 2 *A rule F satisfies UD, CY, RM and TP if and only $F \in \{L_\varphi\}_{\varphi\in\Phi}$.*

Theorem 2, in conjunction with Proposition 3, shows that the axioms UD, CY, RM and TP are equivalent to a kind of egalitarianism, where the equality in question is equality of a conception of well-being that is a general index of welfare and resources. In particular, prioritarianism, at least in conjunction with solidarity, does not preclude equality, but it modifies the equalisandum from 'welfare' to an index of welfare and resources.

The reader may consult the Appendix to verify that all the properties in Theorem 2 are independent. There, we give some examples of rules which satisfy all the properties in the theorem except the one that is explicitly mentioned.

The proof of Theorem 2 also appears in the Appendix. A close examination of this proof shows that it can also be applied to the domain $\mathcal{E}^c$. As a result, and making use of Proposition 1, we have the following characterization.

Theorem 2* *A rule F satisfies RD, CY, RM, SP and LP if and only $F \in \{L_\varphi\}_{\varphi\in\Phi}$.*

It is straightforward to show that the family $\{L_\varphi\}_{\varphi\in\Phi}$ satisfies CY^*, a stronger version of CY in which no restriction on the cardinality of the seceding group is imposed. We show in the Appendix that CY^* and RM together are equivalent to AG. Thus, we have the following results.

Corollary 1 *The family of rules $\{L_\varphi\}_{\varphi\in\Phi}$ is characterized by UD, AG, and TP.*

Corollary 2 *The family of rules* $\{L_\varphi\}_{\varphi\in\Phi}$ *is characterized by RD, AG, SP and LP.*

Note that, as promised earlier, Corollaries 1 and 2 have eliminated the consistency axiom.

We have shown (Proposition 2) that SP is a consequence of HE and AM. It is straightforward to show that the family $\{L_\varphi\}_{\varphi\in\Phi}$ satisfies AM. Since HE is implied by LP, the following characterization is also obtained.[21]

Corollary 3 *The family of rules* $\{L_\varphi\}_{\varphi\in\Phi}$ *is characterized by RD, AG, AM and LP.*

This last result shows that impartiality (RD), solidarity (AG and AM) and limited priority (LP) imply the equalization of an index of resources and welfare.

Corollaries 1 and 3 will motivate for the reader our development of the theory on the two domains $\mathcal{E}$ and $\mathcal{E}^c$. With respect to our stated concern, of deriving prioritarianism from solidarity and impartiality, Corollary 3 (which assumes the restricted domain $\mathcal{E}^c$) is a more satisfying result than Corollary 1. In Corollary 3, we deduce semi-priority from other axioms (granted, one of which is limited priority). On the domain $\mathcal{E}$, however, we are not able to 'factor out' semi-priority. Corollary 1 assumes the strong axiom TP as a premise.

4.3 Two important allocation rules

In this section we focus in two rules within the family of $\{L_\varphi\}_{\varphi\in\Phi}$ rules. The **equal resource** (ER) rule is the L_{φ_1} rule, where $\varphi_1(x,y)=x$. The **equal welfare** (EW) rule is the L_{φ_2} rule, where $\varphi_2(x,y)=y$. The ER rule equalizes wealth in all economies, whereas the EW rule equalizes welfare in all economies. These two rules are the extreme prioritarian rules for the most

[21] One also obtains as immediate consequences that the family of rules $\{L_\varphi\}_{\varphi\in\Phi}$ is characterized by RD, AG, AM, and TP and by RD, CY, RM, AM, and LP.

able and the least able agents in an economy. More precisely, ER is the best (worst) prioritarian rule for the ablest (disablest) agent, whereas EW is the best (worst) prioritarian rule for the disablest (ablest) agent.

Proposition 4 *Let $e = (I, u, W) \in \mathcal{E}^c$. Let i (j) be the ablest (disablest) individual in I. Then, for all rules F satisfying SP and LP we have the following:*
(i) $ER_i(e) \geq F_i(e) \geq EW_i(e)$
(ii) $ER_j(e) \leq F_j(e) \leq EW_j(e)$

Proof.

Let F be a rule satisfying SP and LP. Let $e = (I, u, W) \in \mathcal{E}^c$ and let i (j) be the ablest (disablest) individual in I. We shall show (i). The proof of (ii) is analogous.

Suppose, contrary to the claim, that $ER_i(e) < F_i(e)$. Then, there exists $k \in I$ such that $ER_k(e) > F_k(e)$. Since $ER_k(e) = ER_i(e)$, it follows that $F_i(e) > F_k(e)$. But this contradicts SP, as $u_i \geq u_k$.

Similarly, if $F_i(e) < EW_i(e)$, there exists $k \in I$ such that $EW_k(e) < F_k(e)$. Since u_i and u_k are strictly increasing, it follows that $u_i(F_i(e)) < u_i(EW_i(e)) = u_k(EW_k(e)) < u_k(F_k(e))$. But this contradicts LP, as $u_i \geq u_k$.
■

In particular, Proposition 4 shows that, for all $\varphi \in \Phi$,

$$ER_i(e) \geq (L_\varphi)_i(e) \geq EW_i(e) \text{ and } ER_j(e) \leq (L_\varphi)_j(e) \leq EW_j(e),$$

where i and j are, respectively, the ablest and disablest individuals in e.

We can define a duality relationship between the members of the $\{L_\varphi\}_{\varphi\in\Phi}$ family as follows. For each $\varphi \in \Phi$, let φ^* be defined as $\varphi^*(x, y) = \varphi(y, x)$. Then, $\varphi^* \in \Phi$. We define the **dual** rule of L_φ as L_{φ^*}. L_φ and L_{φ^*} are symmetric with respect to the treatment of wealth and welfare. Note that ER and EW are dual rules.

5 Recapitulation

To review, we have been concerned with the following syllogism:

A. Justice requires impartiality;

B. Impartiality, as far as justice is concerned, is properly modeled by veil-of-ignorance thought experiments;

C. Veil-of-ignorance thought experiments in general recommend anti-prioritarian allocations.

Therefore,

D. Justice is not prioritarian.

We reject D. A has a long intellectual history, and we do not reject it. C is, so far as we can tell, a fact.

Those who reject D can avail themselves of at least the following possible strategies:

1. To construct a model of the VI that does not conflict with prioritarianism, thus negating C. Perhaps this can be done. Our approach has been to formalize two of the best models of the veil of ignorance offered in the last half century and to show they are anti-prioritarian. But this is not a proof that C is true.

2. To refine the definition of impartiality to exclude the veil of ignorance. Perhaps this can be done. We take this to be the strategy of Brian Barry–how else could he claim that justice is (or as) impartiality, and also believe that justice is prioritarian or more? Perhaps this is also Scanlon's (1998) strategy: we leave this for others to judge.

3. To admit that a second principle (after impartiality) is required to characterize justice. We have chosen solidarity–*de gustibus non disputandum est*. One might also profitably employ fraternity or reciprocity.

Strategy 3 appears to succeed, at least in very simple environments. Indeed, we get something more: solidarity and impartiality imply a kind of egalitarianism, where the *index* of wealth and welfare that is equalized according to justice is not determined without further assumptions. Two classical distribution rules are polar (and even 'dual') in the class of index-egalitarian rules–the equal resource and equal welfare allocation rules.

We note that indices of goods of various kinds have often appeared in the recent philosophical literature. Rawls writes that justice requires the maximinization of an index of primary goods. Amartya Sen's (1980, 1992) proposal to equalize capabilities is often formulated in terms of equalizing an index of his 'functionings': indeed, the major practical application of his theory, by the UNDP in its annual *Human Development Report*, computes an index of various functionings for a large set of countries.

We believe ours is the first work that axiomatizes an index of wealth and welfare as the equalisandum of a theory of justice.

We remark upon our specification of the domain of economies, in which the risk preferences of individuals are not described. Obviously, this excludes veil-of-ignorance type allocation rules by fiat, assuming that VI constructions must exploit the distribution of risk preferences of persons in the world, as the Harsanyi and Dworkin veils do.

It would be preferable, of course, to include risk preferences of individuals in the description of economic environments, as well as the profile of level-comparable utility functions. To do so, however, would immensely complicate the analysis, as it would vastly increase the number of allocation rules that can be defined.

Eventually, a theory of distributive justice must, we believe, postulate a domain of economies in which risk preferences, level-comparable welfare, and effort choices by individuals (relating to education and production) are described. The present analysis is a far cry from that goal. Indeed, one difficulty in the work of philosophers is that they implicitly assume all these attributes of real-world societies in their theorizing. Clearly, it would be immensely difficult to deduce formally a theory of just resource allocation on such a domain, without postulating unacceptably strong axioms, and so it is not surprising that the work of political philosophers is tentative and sketchy, by their own admission.

Qualitatively, we have suggested that left-liberal principles of distributive justice cannot be deduced without supposing axioms of a cooperative sort. The hope of deducing, from universally appealing axioms, the degree of equality prized by those on the liberal left is, we believe, empty.

Solidarity, as we have modeled it, is in conflict with self-interest. A self-interested person has no reason to donate to a fellow citizen who has become ill. In a dynamic model, however, where persons' lives extend over many periods, one might hope to show that solidarity is *not* in conflict with self-interest: Andrea contributes to Bob's support when he becomes ill today, and Bob (or his descendent) returns the favor when she (or her descendent) becomes ill in the future. This is reciprocity, à la Serge Kolm (2004). Solidarity, therefore, would be a form of insurance. It could therefore be the case, in a dynamic model, when we construct a veil of ignorance in which worldly persons remain self-interested, that their representative souls adopt solidarity as a principle governing wealth allocation. It could transpire, were that the case, that the conflict between the veil and prioritarianism would vanish, and the present paper would become obsolete–its focus being upon a pathology that only occurs because of the overly simple (static) nature of the model.

We do not conjecture, however, that this happy ending will transpire. For, were it to, it would be because selfish agents viewed solidaristic trade as a good form of insurance. We know, however, from the static case, that self-interested agents do not always insure against bad outcomes–that is a lesson of Sections 2 and 3 of this paper. We believe, likewise, that in a dynamic setting, if agents are not too risk averse, they will not see solidaristic behavior as to their (long-run) advantage. Therefore, we conjecture that, simple as the models of this paper are, the lessons they produce may well endure in the dynamic setting.

A final comment is in order concerning our axiom of priority. What we, the ethical observers, consider a bad outcome–having a society in which the disabled are less abundantly resourced than the able–does not coincide with the bad outcome to the individual behind the veil who faces the birth lottery; the worst outcome for her may be being born *able* without *sufficient* resources to fully exploit that ability. This is, of course, why the veil sometimes (often) allocates less wealth to the disabled than to the able. We have not *defended* our axiom of priority: it is, after all, an axiom. To do so would require a fully philosophical inquiry.

6 Appendix

In this section we provide the proof of Theorem 2. We first develop the necessary machinery for this proof by means of a series of lemmata. Then, we present the proof and finally we show that the result is tight.

6.1 Preliminary results

The first lemma shows that an agent can achieve any level of wealth in a two-agent economy, as long as the allocation rule satisfies RM and TP.

Lemma 1 *Let F be a rule that satisfies UD, RM and TP. Let $i, j \in \mathbb{I}$ be given and $\alpha > 0$. Then, there is a value W such that $F_i(e) = \alpha$, where $e = (\{i, j\}, (u_i, u_j), W) \in \mathcal{E}$.*

Proof.

Let F be a rule that satisfies UD, RM and TP. Let $i, j \in \mathbb{I}$ and $\alpha > 0$ be given and denote $e = (\{i, j\}, (u_i, u_j), W) \in \mathcal{E}$ for each $W > 0$. Since $F_i(e) \leq W$, it follows that, for W sufficiently small, $F_i(e) < \alpha$.

Suppose that $F_i(e) < \alpha$ for all e. Then, $F_j(e) > W - \alpha$ for all e and therefore $\lim_{W\to\infty} F_j(e) = \infty$. In particular, for all e such that $W > 2\alpha$, $F_j(e) > \alpha > F_i(e)$. Thus, by TP, $u_i(F_i(e)) \geq u_j(F_j(e))$. Since u_i is increasing, $u_i(\alpha) \geq u_j(F_j(e))$ for all e such that $W > 2\alpha$. However, since u_j is unbounded, $\lim_{W\to\infty} u_j(F_j(e)) = \infty$, a contradiction.

Thus, there exist W_1 and W_2 such that $F_i(e_1) < \alpha$ and $F_i(e_2) > \alpha$ for $e_1 = (\{i,j\},(u_i,u_j),W_1)$ and $e_2 = (\{i,j\},(u_i,u_j),W_2)$. Consider the following two sets:

$$\Omega^< = \{W \in (0,+\infty) : F_i(e) < \alpha\} \text{ and } \Omega^> = \{W \in (0,+\infty) : F_i(e) > \alpha\}.$$

Then, $W_1 \in \Omega^<$ and $W_2 \in \Omega^>$. Thus,

$$\Omega^< \neq \emptyset \text{ and } \Omega^> \neq \emptyset. \tag{11}$$

Furthermore, it is obvious that

$$\Omega^< \cap \Omega^> = \emptyset. \tag{12}$$

We show now that

$$\Omega^< \text{and } \Omega^> \text{are open sets.} \tag{13}$$

Claim. $\Omega^<$ and $\Omega^>$ are open sets.
Proof of the claim. Let $W \in \Omega^<$ and $\overline{\alpha} = F_i(e) < \alpha$. Let $\varepsilon = \frac{\alpha - \overline{\alpha}}{2}$. We show that $(W-\varepsilon, W+\varepsilon) \subset \Omega^<$. By RM, $(W-\varepsilon, W) \subset \Omega^<$. Suppose, by contradiction, that there exists $W^* \in (W, W+\varepsilon)$ such that $W^* \notin \Omega^<$, i.e., $F_i(e^*) \geq \alpha$, for $e^* = (\{i,j\},(u_i,u_j),W^*)$. By RM, $F_j(e^*) > F_j(e) = W - \overline{\alpha}$. Then, $W^* = F_j(e^*) + F_i(e^*) > W - \overline{\alpha} + \alpha = W + 2\varepsilon$, which contradicts that $W^* \in (W, W+\varepsilon)$. This shows that $\Omega^<$ is an open set.

Let $W \in \Omega^>$ and $\overline{\alpha} = F_i(e) > \alpha$. Let $\varepsilon = \frac{\overline{\alpha} - \alpha}{2}$. We show that $(W - \varepsilon, W + \varepsilon) \subset \Omega^>$. By RM, $(W, W+\varepsilon) \subset \Omega^>$. Suppose, by contradiction, that there exists $W^* \in (W-\varepsilon, W)$ such that $W^* \notin \Omega^>$, i.e., $F_i(e^*) \leq \alpha$, for $e^* = (\{i,j\},(u_i,u_j),W^*)$. By RM, $F_j(e^*) < F_j(e) = W - \overline{\alpha}$. Then, $W^* = F_j(e^*) + F_i(e^*) < W - \overline{\alpha} + \alpha = W - 2\varepsilon$, which contradicts that $W^* \in (W-\varepsilon, W)$. This shows that $\Omega^>$ is an open set, which proves the claim.

Now, if, contrary to the statement of the lemma, $F_i(e) \neq \alpha$, for all $W \in \mathbb{R}_{++}$, then

$$\mathbb{R}_{++} \subset \Omega^> \cup \Omega^<. \tag{14}$$

Finally, (11), (12), (13) and (14) together say that $(0, +\infty)$ is not connected, which is a contradiction. ■

It is worth noting that, by Proposition 1, SP and LP together are equivalent to TP, in $\mathcal{E}^c$. Since no argument in the above proof uses that agents are not comparable, we also obtain that an agent can achieve any level of wealth in a two-agent economy, when the domain is $\mathcal{E}^c$ and the allocation rule satisfies RD, RM, SP and LP. Formally,

Lemma 1* *Let F be a rule that satisfies RD, RM, SP and LP. Let $i, j \in \mathbb{I}$ be two given comparable agents and $\alpha > 0$. Then, there is a value W such that $F_i(e) = \alpha$, where $e = (\{i, j\}, (u_i, u_j), W) \in \mathcal{E}^c$.*

Now, for an allocation rule F, an individual $i \in \mathbb{I}$, and $\alpha \in \mathbb{R}_+$, we define the domain of economies for which individual i obtains an amount of wealth α. Formally:

$$E(F, i, \alpha) = \{e = (I, u, W) \in \mathcal{E} : i \in I \text{ and } F_i(e) = \alpha\}.$$

It is straightforward to show that if a rule satisfies strong priority, everyone receives at least some wealth. Thus, if F satisfies TP the only economies in $E(F, i, 0)$ are those with $W = 0$.

We also define the sets

$$C(F, i, \alpha) = \{(a, b) \in \mathbb{R}_+^2 : \exists e = (I, u, W) \in E(F, i, \alpha),\ j \in I \text{ s.t. } (a, b) = (F_j(e), u_j(F_j(e)))\}.$$

In words, $C(F, i, \alpha)$ is the set of points in the plane which are achieved as wealth-welfare ordered pairs under the action of F on individuals who are members of economies in $E(F, i, \alpha)$. As before, if F satisfies TP then $C(F, i, 0) = \{(0, 0)\}$ for all $i \in \mathbb{I}$, whereas $(0, 0) \notin C(F, i, \alpha)$ for any $i \in \mathbb{I}$ and $\alpha > 0$. As a consequence of Lemma 1, we have that, given $C(F, i, \alpha)$ and $j \in \mathbb{I}$, there exists $(a, b) \in C(F, i, \alpha)$ such that $b = u_j(a)$. Finally, it is also straightforward to show that, if F satisfies CY, then

$$C(F, i, \alpha) = \{(a, b) \in \mathbb{R}_+^2 : \exists j \in \mathbb{I} \text{ s.t. } (a, b) = (F_j(e), u_j(a)), \text{ where } e = (\{i, j\}, (u_i, u_j), a+\alpha)\}.$$

We can define in an analogous way the sets $E^c(F, i, \alpha)$ and $C^c(F, i, \alpha)$ corresponding to the domain of economies with comparable individuals. If F satisfies RD, SP and LP then $C^c(F, i, 0) = \{(0, 0)\}$ for all $i \in \mathbb{I}$, whereas $(0, 0) \notin C^c(F, i, \alpha)$ for any $i \in \mathbb{I}$ and $\alpha > 0$.

Our aim is to show that for fixed F and i, and varying α the family of curves $\{C(F, i, \alpha)\}$, or alternatively $\{C^c(F, i, \alpha)\}$, is the isoquant map of an appropriate increasing, continuous function φ. We first have the following.

Lemma 2 *Let F be a rule that satisfies UD, RM and TP. Let $i \in \mathbb{I}$ and $(a, b) \in \mathbb{R}^2_+$. Then, there exists $\alpha \in \mathbb{R}_+$ such that $(a, b) \in C(F, i, \alpha)$.*

Proof.

Let F be a rule that satisfies UD, RM and TP. Let $(a, b) \in \mathbb{R}^2_+$ be given. Take an agent $j \in \mathbb{I}$ for whom $u_j(a) = b$. By Lemma 1 there exists $W \in \mathbb{R}_+$ such that $F_j(e) = a$, where $e = (\{i, j\}, (u_i, u_j), W)$. Let $\alpha = W - a$. Then, $(a, b) \in C(F, i, \alpha)$. ■

Note that the proof of Lemma 2 is also valid to show the following:

Lemma 2* *Let F be a rule that satisfies RD, RM, SP and LP. Let $i \in \mathbb{I}$ and $(a, b) \in \mathbb{R}^2_+$. Then, there exists $\alpha \in \mathbb{R}_+$ such that $(a, b) \in C^c(F, i, \alpha)$.*

The next lemma says if F satisfies UD, CY and RM then agents belonging to more than one economy in $E(F, i, \alpha)$ will be allocated the same level of resource in all these economies.

Lemma 3 *Let F be a rule that satisfies UD, CY and RM. Let $i \in \mathbb{I}$ and $\alpha \in \mathbb{R}_+$. Then, for all $e = (I, u, W) \in E(F, i, \alpha)$ and $e' = (I', u', W') \in E(F, i, \alpha)$ we have $F_j(e) = F_j(e')$ for all $j \in I \cap I'$.*

Proof.

Suppose the claim were false; then there is a pair of economies $e = (I, u, W)$ and $e' = (I', u', W') \in E(F, i, \alpha)$ such that $j \in I \cap I'$ and $W_j = F_j(e) > F_j(e') = W'_j$. Consider the economies

$$\widehat{e} = (\{i, j\}, (u_i, u_j), \alpha + W_j) \text{ and } \widehat{e}' = (\{i, j\}, (u_i, u_j), \alpha + W'_j).$$

By RM, $F_i(\widehat{e}) > F_i(\widehat{e}')$. By CY, however, $F_i(\widehat{e}) = \alpha = F_i(\widehat{e}')$. ■

The above proof suffices to show the equivalent result in $\mathcal{E}^c$.

Lemma 3* *Let F be a rule that satisfies RD, CY and RM. Let $i \in \mathbb{I}$ and $\alpha \in \mathbb{R}_+$. Then, for all $e = (I, u, W) \in E^c(F, i, \alpha)$ and $e' = (I', u', W') \in E^c(F, i, \alpha)$ we have $F_j(e) = F_j(e')$ for all $j \in I \cap I'$.*

Now we show that if F satisfies CY, RM and TP, then the curve generated by $C(F, i, \alpha)$ slopes down to the right.[22]

Lemma 4 *Let F be a rule that satisfies UD, CY, RM and TP. Let $i \in \mathbb{I}$ and $\alpha \in \mathbb{R}_+$. If $(a, b) \in C(F, i, \alpha)$, $(a', b') \in C(F, i, \alpha)$ and $a' > a$ then $b' \leq b$.*

Proof.

Suppose, to the contrary, that $b' > b$. By definition, there exist $e = (I, u, W)$ and $e' = (I', u', W') \in E(F, i, \alpha)$ and $j \in I$, $k \in I'$ such that $(a, b) = (F_j(e), u_j(F_j(e)))$, $(a', b') = (F_k(e'), u_k(F_k(e')))$. As well, there is a wealth W^* such that $e^* = (\{i, j, k\}, (u_i, u_j, u_k), W^*) \in E(F, i, \alpha)$ (same argument as Lemma 1). By Lemma 3, we know that $F_j(e^*) = a$, $F_k(e^*) = a'$ and so $F_j(e^*) < F_k(e^*)$. Thus, by TP, $u_j(F_j(e^*)) \geq u_k(F_k(e^*))$. However, we also know that, by hypothesis, $b = u_j(F_j(e^*)) < u_k(F_k(e^*)) = b'$. This contradiction establishes the lemma. ■

Analogously,

Lemma 4* *Let F be a rule that satisfies RD, CY, RM, SP and LP. Let $i \in \mathbb{I}$ and $\alpha \in \mathbb{R}_+$. If $(a, b) \in C^c(F, i, \alpha)$, $(a', b') \in C^c(F, i, \alpha)$ and $a' > a$ then $b' \leq b$.*

We define $\Lambda(C(F, i, \alpha))$, the **support** of the curve $C(F, i, \alpha)$, as the wealth values for which there exist welfare levels such that the pairs are achieved

[22] The fact that $C(F, i, \alpha)$ is a curve is indeed a consequence of Lemma 4. This lemma shows, in particular, that the interior of the set $C(F, i, \alpha)$ is empty. In other words, no ball centered in a point of $C(F, i, \alpha)$ is completely included in that set.

under the action of F. Formally:

$$\begin{aligned}\Lambda(C(F,i,\alpha)) &= \{a \in \mathbb{R}_+ : \exists\, b \in \mathbb{R}_+ \text{ s.t. } (a,b) \in C(F,i,\alpha)\} \\ &= \{a \in \mathbb{R}_+ : \exists\, e = (I,u,W) \in E(F,i,\alpha) \text{ s.t. } a = F_j(e) \text{ for some } j \in I\}.\end{aligned}$$

We define $\Gamma(C(F,i,\alpha))$, the **image** of the curve $C(F,i,\alpha)$, as the welfare levels for which there exist wealth values such that the pairs are achieved under the action of F. Formally,

$$\begin{aligned}\Gamma(C(F,i,\alpha)) &= \{b \in \mathbb{R}_+ : \exists\, a \in \mathbb{R}_+ \text{ s.t. } (a,b) \in C(F,i,\alpha)\} \\ &= \{b \in \mathbb{R}_+ : \exists\, e = (I,u,W) \in E(F,i,\alpha) \text{ s.t. } b = u_j(F_j(e)) \text{ for some } j \in I\}.\end{aligned}$$

Another property is obtained:

Lemma 5 *Let F be a rule that satisfies UD, CY, RM and TP. Let $i \in \mathbb{I}$ and $\alpha \in \mathbb{R}_+$. Then, $C(F,i,\alpha)$ is a connected set.*

Proof.

Let F be a rule that satisfies UD, CY, RM and TP. Fix $i \in \mathbb{I}$ and $\alpha \in \mathbb{R}_+$. We define $\gamma : \Lambda(C(F,i,\alpha)) \to \Gamma(C(F,i,\alpha))$, the mapping whose graph coincides with the curve $C(F,i,\alpha)$, i.e., $Gr(\gamma) = C(F,i,\alpha)$.[23]

Suppose, contrary to the claim, that the curve $C(F,i,\alpha)$ is not connected. Since $C(F,i,\alpha)$ is downward sloping, then either $\Lambda(C(F,i,\alpha))$ is not connected or $\Gamma(C(F,i,\alpha))$ is not connected.

Case 1: $\Lambda(C(F,i,\alpha))$ is not connected.

Let $a,b \in \Lambda(C(F,i,\alpha))$ such that $a < b$ and $\lambda \cdot a + (1-\lambda)b \notin \Lambda(C(F,i,\alpha))$ for all $\lambda \in (0,1)$. Since $C(F,i,\alpha)$ is downward sloping, it follows that $\min\{\gamma(a)\} \geq \max\{\gamma(b)\}$. Fix $\widehat{\lambda} \in (0,1)$ and let $\widehat{x} = \widehat{\lambda} \cdot a + (1-\widehat{\lambda})b$. Consider $\theta = \frac{\min\{\gamma(a)\}}{\overline{x}}$ and let $u_j(x) = \theta \cdot x$, for all $x \in \mathbb{R}_+$. Then, $u_j \in \mathcal{U}$ and

$$\max\{\gamma(b)\} \leq u_j(\widehat{x}) = \min\{\gamma(a)\}.$$

[23] Note that γ may well be a multi-valued function. For generality, the proof will be framed assuming that γ is multi-valued.

We know, by Lemma 1, that there exists $w \in \mathbb{R}_+$ such that $(w, u_j(w)) \in C(F, i, \alpha)$. Thus, $u_j(w) \cap \gamma(w) \neq \emptyset$.

Now, since $C(F, i, \alpha)$ is downward sloping, we have that

$$y < u_j(\widehat{x}) \text{ for all } y \in \gamma(x) \text{ such that } x \in \Lambda(C(F, i, \alpha)) \text{ and } x > b,$$

and

$$y > u_j(\widehat{x}) \text{ for all } y \in \gamma(x) \text{ such that } x \in \Lambda(C(F, i, \alpha)) \text{ and } x < a.$$

Since u_j is strictly increasing, it follows that $u_j(x)$ and $\gamma(x)$ do not cross, which is a contradiction.

Insert **Figure 1** about here

Case 2: $\Gamma(C(F, i, \alpha))$ is not connected.

Let $a, b \in \Gamma(C(F, i, \alpha))$ such that $a < b$ and $\lambda \cdot a + (1-\lambda)b \notin \Gamma(C(F, i, \alpha))$ for all $\lambda \in (0, 1)$. Assume there exists $\overline{x}$ such that $\{a, b\} \subset \gamma(\overline{x})$.[24] Fix $\widehat{\lambda} \in (0, 1)$ and let $\theta = \frac{\widehat{\lambda} \cdot a + (1-\widehat{\lambda})b}{\overline{x}}$. Consider $u_j(x) = \theta \cdot x$, for all $x \in \mathbb{R}_+$. Then, $u_j \in \mathcal{U}$ and $u_j(\widehat{x}) = \widehat{\lambda} \cdot a + (1-\widehat{\lambda})b$. We know, by Lemma 1, that there exists $w \in \mathbb{R}_+$ such that $(w, u_j(w)) \in C(F, i, \alpha)$. Thus, $u_j(w) \cap \gamma(w) \neq \emptyset$.

Now, since $C(F, i, \alpha)$ is downward sloping, it follows that

$$y < u_j(\widehat{x}) \text{ for all } y \in \gamma(x) \text{ such that } x \in \Lambda(C(F, i, \alpha)) \text{ and } x > \widehat{x},$$

and

$$y > u_j(\widehat{x}) \text{ for all } y \in \gamma(x) \text{ such that } x \in \Lambda(C(F, i, \alpha)) \text{ and } x < \widehat{x}.$$

Since u_j is strictly increasing, it follows that $u_j(x)$ and $\gamma(x)$ do not cross, which is a contradiction. ■

Note that we have shown, in particular, that both $\Lambda(C(F, i, \alpha))$ and $\Gamma(C(F, i, \alpha))$ are connected.

[24] If this were not the case, then $\Lambda(C(F, i, \alpha))$ would not be connected, and the proof of Case 1 would be valid to conclude.

A close examination of the above proof shows that it can be extended to the domain $\mathcal{E}^c$:

Lemma 5* *Let F be a rule that satisfies RD, CY, RM, SP and LP. Let $i \in \mathbb{I}$ and $\alpha \in \mathbb{R}_+$. Then, $C^c(F, i, \alpha)$ is a connected set.*

Now, we show that the sets $\{C(F, i, \alpha)\}$ are disjoint.

Lemma 6 *Let F be a rule that satisfies UD, CY, RM and TP. Let $i \in \mathbb{I}$, $\alpha_1 \in \mathbb{R}_+$ and $\alpha_2 \in \mathbb{R}_+$ such that $\alpha_1 \neq \alpha_2$. Then, $C(F, i, \alpha_1) \cap C(F, i, \alpha_2) = \emptyset$.*

Proof.

Let $\alpha_1 > \alpha_2$. Suppose $(a, b) \in C(F, i, \alpha_1) \cap C(F, i, \alpha_2)$. Let $e_1 = (I_1, u_1, \alpha_1) \in E(F, i, \alpha_1)$, $e_2 = (I_2, u_2, \alpha_2) \in E(F, i, \alpha_2)$ and $j \in I_1$, $k \in I_2$ such that $(a, b) = (F_j(e_1), u_j(F_j(e_1))) = (F_k(e_2), u_k(F_k(e_2)))$. Consider the economies $\widehat{e}_1 = (\{i, j\}, (u_i, u_j), a + \alpha_1)$ and $\widehat{e}_2 = (\{i, k\}, (u_i, u_k), a + \alpha_2)$. CY implies that $F_i(\widehat{e}_1) = \alpha_1$ and $F_i(\widehat{e}_2) = \alpha_2$. By a similar argument to that of Lemma 1, there is a $W > a + \alpha_2$ such that $\widetilde{e}_2 = (\{i, k\}, (u_i, u_k), W) \in E(F, i, \alpha_1)$. By RM, applied to $\widehat{e}_2$ and $\widetilde{e}_2$ we know that $F_k(\widetilde{e}_2) > F_k(\widehat{e}_2) = a$. Therefore, $(a, b) < (F_k(\widetilde{e}_2), u_k(F_k(\widetilde{e}_2))) \in C(F, i, \alpha_1)$. This contradicts the fact that $C(F, i, \alpha_1)$ is downward sloping, as Lemma 4 shows. ■

Analogously,

Lemma 6* *Let F be a rule that satisfies RD, CY, RM, SP and LP. Let $i \in \mathbb{I}$, $\alpha_1 \in \mathbb{R}_+$ and $\alpha_2 \in \mathbb{R}_+$ such that $\alpha_1 \neq \alpha_2$. Then, $C^c(F, i, \alpha_1) \cap C^c(F, i, \alpha_2) = \emptyset$.*

The next lemma requires a preliminary definition. Of two sets A and B in the plane we say that B lies above A if

1. For all $(a_1, a_2) \in A$ there exists $(b_1, b_2) \in B$ such that $(a_1, a_2) < (b_1, b_2)$.
2. There is no $(a_1, a_2) \in A$ and $(b_1, b_2) \in B$ such that $(b_1, b_2) < (a_1, a_2)$.

Clearly, if B lies above A, then A does not lie above B.

Lemma 7 *Let F be a rule that satisfies UD, CY, RM and TP. Let $i \in \mathbb{I}$, $\alpha_1 \in \mathbb{R}_+$ and $\alpha_2 \in \mathbb{R}_+$ such that $\alpha_1 > \alpha_2$. Then, $C(F, i, \alpha_1)$ lies above $C(F, i, \alpha_2)$.*

Proof.

Let $(a, b) \in C(F, i, \alpha_2)$, and let $e = (\{i, j\}, (u_i, u_j), W) \in E(F, i, \alpha_2)$ such that $F_j(e) = a$, $u_j(F_j(e)) = b$. Since $F_i(e) = \alpha_2$, and by a similar argument to that of the proof of Lemma 1, increasing the wealth from its value in e, we eventually find a wealth value W^* such that $e^* = (\{i, j\}, (u_i, u_j), W^*) \in E(F, i, \alpha_1)$. Let $(a', b') = (F_j(e^*), u_j(F_j(e^*)))$. Then, $(a', b') \in C(F, i, \alpha_1)$. Furthermore, since F satisfies RM and u_j is strictly increasing, $(a', b') > (a, b)$.

Conversely, let $(a, b) \in C(F, i, \alpha_1)$. Suppose there were a point $(a'', b'') \in C(F, i, \alpha_2)$ such that $(a'', b'') > (a, b)$. We know that there is a point (α_1, d_1) in $C(F, i, \alpha_1)$–because $(F_i(e), u_i(F_i(e))) = (\alpha_1, d_1)$ for any e in $E(F, i, \alpha_1)$– and in like manner there is a point $(\alpha_2, d_2) \in C(F, i, \alpha_2)$, and $d_1 = u_i(\alpha_1) > u_i(\alpha_2) = d_2$ (because both points are associated with agent i). Thus we have

$$\begin{aligned} C(F, i, \alpha_1) &\ni (\alpha_1, d_1) > (\alpha_2, d_2) \in C(F, i, \alpha_2), \\ C(F, i, \alpha_2) &\ni (a'', b'') > (a, b) \in C(F, i, \alpha_1) \end{aligned} \tag{15}$$

Without loss of generality, assume that $\alpha_2 < a''$. Then, it follows that $\alpha_2 < \alpha_1 < a < a''$.[25]

Insert **Figure 2** about here

For $k = 1, 2$, let $\gamma^k : \Lambda(C(F, i, \alpha_k)) \to \mathbb{R}_+$ be the mapping whose graph is $C(F, i, \alpha_k)$. Then, by Lemma 4 and (15) we have

$$\begin{aligned} &\max\{\gamma^2(\alpha_1)\} \leq \min\{\gamma^2(\alpha_2)\} \leq \max\{\gamma^2(\alpha_2)\} < \min\{\gamma^1(\alpha_1)\} \text{ and} \\ &\min\{\gamma^2(a)\} \geq \max\{\gamma^2(a'')\} \geq \min\{\gamma^2(a'')\} > \max\{\gamma^1(a)\}. \end{aligned} \tag{16}$$

By Lemma 5, both $Gr(\gamma^1) = C(F, i, \alpha_1)$ and $Gr(\gamma^2) = C(F, i, \alpha_2)$ are connected sets. So are their supports and therefore

$$(\alpha_1, a) \subset \Lambda(C(F, i, \alpha_1)) \cap \Lambda(C(F, i, \alpha_2)).$$

[25] If $\alpha_2 > a''$, then $a < a'' < \alpha_2 < \alpha_1$, and the ensuing argument would be analogous.

Since $\max\{\gamma^2(\alpha_1)\} < \min\{\gamma^1(\alpha_1)\}$ and $\min\{\gamma^2(a)\} > \max\{\gamma^1(a)\}$, it follows that there exists $x \in (\alpha_1, a)$ such that $\gamma^1(x) \cap \gamma^2(x) \neq \emptyset$, which means that $C(F, i, \alpha_1) \cap C(F, i, \alpha_2) \neq \emptyset$. This contradicts Lemma 6. ■

In an analogous way, we can prove the following result:

Lemma 7* *Let F be a rule that satisfies RD, CY, RM, SP and LP. Let $i \in \mathbb{I}$, $\alpha_1 \in \mathbb{R}_+$ and $\alpha_2 \in \mathbb{R}_+$ such that $\alpha_1 > \alpha_2$. Then, $C^c(F, i, \alpha_1)$ lies above $C^c(F, i, \alpha_2)$.*

6.2 Proof of the characterization result

We are now ready to prove Theorem 2.

Theorem 2 *A rule F satisfies UD, CY, RM and TP if and only $F \in \{L_\varphi\}_{\varphi \in \Phi}$.*

Proof.

Let us see first that an L_φ rule satisfies all the mentioned properties.

It satisfies CY trivially: if in the sub-economy wealth could be reallocated to increase the leximin order on the pair of agents, then the original allocation would not have been leximinned.

For all $i \in \mathbb{I}$, define $\psi_i : \mathbb{R}_+ \to \mathbb{R}_+$ where $\psi_i(x) = \varphi(x, u_i(x))$. Since $\varphi \in \Phi$, it follows that all the ψ_i functions are strictly increasing. Thus, L_φ satisfies RM.

Finally, we show that L_φ satisfies TP. Let $e = (I, u, W) \in \mathcal{E}$ and $i, j \in I$ such that $(L_\varphi)_i(e) < (L_\varphi)_j(e)$. Let $w_i = (L_\varphi)_i(e)$ and $w_j = (L_\varphi)_j(e)$. Suppose $u_i(w_i) < u_j(w_j)$; then $\varphi(w_i, u_i(w_i)) < \varphi(w_j, u_j(w_j))$, which contradicts the fact established in Proposition 3, that the φ index is equalized for all agents under L_φ. Therefore, $u_i(w_i) \geq u_j(w_j)$, which proves that TP holds.

Now, let F be a rule that satisfies UD, CY, RM and TP. Let $i \in \mathbb{I}$ be given. By Lemma 2, for each $(a, b) \in \mathbb{R}^2_+$ there exists $\alpha \in \mathbb{R}_+$ such that $(a, b) \in C(F, i, \alpha)$. By Lemma 6, α is unique. Define then the function

$\varphi^i : \mathbb{R}^2_{++} \cup \{(0,0)\} \to \mathbb{R}_+$ by $\varphi^i(a,b) = \alpha$, where $\alpha \in \mathbb{R}_+$ is the unique number for which $(a,b) \in C(F,i,\alpha)$.

Claim Let φ^i be defined as above. Then, $\varphi^i \in \Phi$.

Proof of the claim. We show first that φ^i has the required monotonicity properties.

By the reasoning in the paragraph preceding Lemma 2, $\varphi^i(0,0) = 0$ and $\varphi^i(x,y) \geq 0$ for all $(x,y) \in \mathbb{R}^2_+$.

Let $x, x', y \in \mathbb{R}_{++}$ such that $x < x'$. If $\varphi^i(x,y) > \varphi^i(x',y)$ then, by Lemma 7, $C(F,i,\varphi^i(x,y))$ lies above $C(F,i,\varphi^i(x',y))$. In such a case, since $(x',y) \in C(F,i,\varphi^i(x',y))$, there exists $(z,t) \in C(F,i,\varphi^i(x,y))$ such that $(x',y) < (z,t)$. Then, $(z,t) > (x,y)$. This contradicts Lemma 4. Thus, $\varphi^i(x,y) \leq \varphi^i(x',y)$. Similarly, for $x,y,y' \in \mathbb{R}_{++}$ such that $y < y'$, $\varphi^i(x,y) \leq \varphi^i(x',y)$.

Finally, let $(x,y),(z,t) \in \mathbb{R}^2_+$ such that $(x,y) > (z,t)$. By Lemma 4 $\varphi^i(x,y) \neq \varphi^i(z,t)$. If $\varphi^i(x,y) < \varphi^i(z,t)$ then, by Lemma 7, $C(F,i,\varphi^i(z,t))$ lies above $C(F,i,\varphi^i(x,y))$. We have, however, that

$$(x,y) \in C(F,i,\varphi^i(x,y)),\ (z,t) \in C(F,i,\varphi^i(z,t)) \text{ and } (x,y) > (z,t).$$

This is a contradiction.

We show now that φ^i is continuous on $\mathbb{R}^2_+$.

Let $\{(a_n,b_n)\}_n$ be a sequence in $\mathbb{R}^2_+$ converging to $(a,b) \in \mathbb{R}^2_+$. We must show that $\{\alpha_n\}_n = \{\varphi(a_n,b_n)\}_n$ converges to $\alpha = \varphi(a,b)$. Of the three sets: $\Omega^> = \{n \in \mathbb{N} : \alpha_n > \alpha\}$, $\Omega^< = \{n \in \mathbb{N} : \alpha_n < \alpha\}$, $\Omega^= = \{n \in \mathbb{N} : \alpha_n = \alpha\}$, at least one is infinite. If $\Omega^=$ is the only infinite set, then the claim is obviously true. So suppose this is not the case; let $\Omega^<$ be infinite. (The proof if $\Omega^>$ is infinite is the same.)

The claim is true unless $\Omega^<$ has a limit point $\overline{\alpha} < \alpha$. Therefore, suppose that this were the case. Denote by $\{\alpha_k\}$ a subsequence of $\Omega^<$ that converges to $\overline{\alpha}$. Consider the curve $C(F,i,\frac{\alpha+\overline{\alpha}}{2})$. There is a ball, B, about $(a,b) \in C(F,i,\alpha)$ which, by Lemma 7, lies above this curve because $\alpha > \frac{\alpha+\overline{\alpha}}{2}$. But for large k, $(a_{\alpha_k}, b_{\alpha_k}) \in B$. This is impossible, since for large k, all points in $C(F, i, \alpha_k)$ lie below $C(F, i, \frac{\alpha+\overline{\alpha}}{2})$. This proves the claim.

Thus, we have shown that for a fixed $i \in \mathbb{I}$ and a rule F satisfying UD, CY, RM and TP, there exists $\varphi^i \in \Phi$ such that $(a, b) \in C(F, i, \varphi^i(a, b))$, for all $(a, b) \in \mathbb{R}^2_+$. We show now that $F = L_{\varphi^i}$.

Fix $e = (I, u, W) \in \mathcal{E}$. Two cases are distinguished.

Case 1: $i \in I$

Let $w_i = F_i(e)$. By definition, $e \in E(F, i, w_i)$ and $(F_j(e), u_j(F_j(e))) \in C(F, i, w_i)$ for all $j \in I$. Therefore, $\varphi^i(F_j(e), u_j(F_j(e))) = w_i$ for all $j \in I$. Then, $F(e) = L_{\varphi^i}(e)$. If the allocation $F(e)$ were not the φ^i-maximin allocation, i.e., if there were an allocation yielding a higher level of φ^i for some agent without lowering the φ^i-values of the others, then the budget constraint would be violated. So on the sub-domain of economies containing agent i, F coincides with the L_{φ^i} rule.

Case 2: $i \notin I$

Pick two agents $j, k \in I$. Let $w_j = F_j(e)$ and $w_k = F_k(e)$. By Lemma 1, there are two economies $\widehat{e} = (\{i, j\}, (u_i, u_j), \widehat{W})$ and $\widetilde{e} = (\{i, k\}, (u_i, u_k), \widetilde{W})$ such that $w_j = F_j(\widehat{e})$ and $w_k = F_k(\widetilde{e})$; let $\widehat{w}_i = F_i(\widehat{e})$ and $\widetilde{w}_i = F_i(\widetilde{e})$.

Claim. $C(F, j, w_j) = C(F, i, \widehat{w}_i)$ and $C(F, k, w_k) = C(F, i, \widetilde{w}_i)$.

Proof of the claim. We only show that $C(F, j, w_j) = C(F, i, \widehat{w}_i)$. The proof of $C(F, k, w_k) = C(F, i, \widetilde{w}_i)$ is identical.

Let $(a, b) \in C(F, i, \widehat{w}_i)$. Then, there exists $l \in \mathbb{I}$ such that $b = u_l(a)$ and $(\widehat{w}_i, a) = (F_i(e^2), F_l(e^2))$, where $e^2 = (\{i, l\}, (u_i, u_l), \widehat{w}_i + a)$. By a similar argument to that of Lemma 1, there exists W^3 such that $F_i(e^3) = \widehat{w}_i$, where $e^3 = (\{i, j, l\}, (u_i, u_j, u_l), W^3)$. Then, $\widehat{e}, e^2$ and e^3 belong to $E(F, i, \widehat{w}_i)$. Thus, by Lemma 3, $a = F_l(e^2) = F_l(e^3)$ and $w_j = F_j(\widehat{e}) = F_j(e^3)$. Consequently, $e^3 \in E(F, j, w_j)$ and $(a, b) \in C(F, j, w_j)$, showing that $C(F, i, \widehat{w}_i) \subseteq C(F, j, w_j)$.

Let $(a, b) \in C(F, j, w_j)$. Then, there exists $l \in \mathbb{I}$ such that $b = u_l(a)$ and $(w_j, a) = (F_j(e^2), F_l(e^2))$, where $e^2 = (\{j, l\}, (u_i, u_l), w_j + a)$. By a similar argument to that of Lemma 1, there exists W^3 such that $F_j(e^3) = w_j$, where $e^3 = (\{i,j,l\}, (u_i, u_j, u_l), W^3)$. Then, $\widehat{e}, e^2$ and e^3 belong to $E(F,j,w_j)$. Thus, by Lemma 3, $a = F_l(e^2) = F_l(e^3)$ and $\widehat{w}_i = F_i(\widehat{e}) = F_i(e^3)$. Consequently, $e^3 \in E(F,i,\widehat{w}_i)$ and $(a,b) \in C(F,i,\widehat{w}_i)$, showing that $C(F,i,\widehat{w}_i) \supseteq C(F,j,w_j)$. This proves the claim.

Note that $(w_j, u_j(w_j)) \in C(F,j,w_j) \cap C(F,k,w_k)$. Since $C(F,j,w_j) = C(F,i,\widehat{w}_i)$ and $C(F,k,w_k) = C(F,i,\widetilde{w}_i)$, then

$$(w_j, u_j(w_j)) \in C(F,i,\widehat{w}_i) \cap C(F,i,\widetilde{w}_i).$$

By Lemma 6, it follows that $\widehat{w}_i = \widetilde{w}_i = w_i$. Thus, $C(F,j,w_j) = C(F,i,w_i) = C(F,k,w_k)$. Therefore all the points $\{(F_l(e), u_l(F_l(e))) : l \in I\}$ lie on the w_i-isoquant of φ^i, and it follows, as in Case 1, that F coincides with L_{φ^i} on the entire domain $\mathcal{E}$. ■

A close examination of this proof shows that it can be applied to the domain $\mathcal{E}^c$, making use of Lemmas 1^*– 7^*. As a result, we have also proven Theorem 2^*.

6.3 On the tightness of the characterization result

In this section, we give examples of rules outside the $\{L_\varphi\}_{\varphi\in\Phi}$ family satisfying all but one of the properties mentioned in Theorem 2. We mention in each case the property that is not fulfilled.

6.3.1 Consistency

Let C be the rule that coincides with the EW rule for two-agent economies and with the ER rule otherwise. Formally, for all $e = (I,u,W) \in \mathcal{E}$,

$$C(e) = \begin{cases} EW(e) \text{ if } |I| = 2 \\ ER(e) \text{ if } |I| > 3 \end{cases}$$

Since both EW and ER satisfy RM and TP, so does C. It fails, however, to satisfy CY. To show this, for all $x \in \mathbb{R}_+$, let $u_1(x) = x$, $u_2(x) = 2x$ and

$u_3(x) = 3x$. Consider $e = (\{1, 2, 3\}, \{u_1, u_2, u_3\}, 3)$. Then, $C(e) = (1, 1, 1)$. Now, let $\overline{e} = (\{1, 2\}, \{u_1, u_2\}, 2)$. Then, $C(\overline{e}) = (\frac{2}{3}, \frac{1}{3}) \neq (1, 1)$.

6.3.2 Resource Monotonicity

Recall that $\sigma_i(W)$ is the wealth that i must receive to enjoy the same level of welfare as individual 0 enjoys with wealth W. We refer hereafter to $\sigma_i(1)$ as the claim of agent i. We define the *constrained equal resource rule* (CER) as the rule that assigns equal resources to all agents in the economy, provided no one receives more than her claim. Obviously, this rule is defined only when the available wealth is lower than the aggregate claim, i.e., $W \leq \sum_{i \in I} \sigma_i(1)$. We extend this rule to the case in which $W > \sum_{i \in I} \sigma_i(1)$ in the following way. Assume, without loss of generality, that $I = \{1, 2, ..., n\}$ and $\sigma_i(1) \leq \sigma_{i+1}(1)$ for all $i = 1, ..., n-1$. As the wealth increases from $\sum_{i \in I} \sigma_i(1)$ to $\sigma_2(1) + \sum_{i=2}^{n} \sigma_i(1)$, then all agents except 1 continue receiving their claims and the remainder is given to agent 1, i.e., $CER(e) = (W - \sum_{i=2}^{n} \sigma_i(1), \sigma_2(1), ..., \sigma_{n-1}(1), \sigma_n(1))$. In general, when the wealth increases from $(k-1) \cdot \sigma_{k-1}(1) + \sum_{i=k}^{n} \sigma_i(1)$ to $k \cdot \sigma_k(1) + \sum_{i=k+1}^{n} \sigma_i(1)$, the $n-k$ agents with the higher claims receive their claims, and the remainder is divided equally among the other agents, i.e., $CER(e) = (\lambda, ..., \lambda, \sigma_{k+1}(1), ..., \sigma_n(1))$, where $\lambda = \frac{W - \left(\sum_{i=k+1}^{n} \sigma_i(1)\right)}{k}$.

Formally, for all $e = (I, u, W) \in \mathcal{E}$,

$$CER(e) = \begin{cases} (\min\{\sigma_i(1), \lambda\})_{i \in I} \text{ if } W \leq \sum_{i \in I} \sigma_i(1) \\ (\max\{\sigma_i(1), \lambda\})_{i \in I} \text{ if } W > \sum_{i \in I} \sigma_i(1) \end{cases},$$

where $\lambda \in \mathbb{R}_+$ is such that $\sum_{i \in I} CER_i(e) = W$.

Proposition 5 *CER satisfies TP and CY but it fails to satisfy RM.*

Proof.

- CER satisfies TP

Let $e = (I, u, W) \in \mathcal{E}$. Let $i, j \in I$ such that $CER_i(e) < CER_j(e)$. We distinguish two cases.

Case 1: $W \leq \sum_{i \in I} \sigma_i(1)$

In this case, $CER_i(e) < CER_j(e)$ if and only if $CER_i(e) = \sigma_i(1)$ and $CER_j(e) = \min\{\sigma_j(1), \lambda\} > \sigma_i(1)$, where λ is such that $\sum_{l \in I} \min\{\sigma_l(1), \lambda\} = W$. If $\sigma_j(1) \leq \lambda$, then $u_j(CER_j(e)) = 1 = u_i(CER_i(e))$. If $\sigma_j(1) > \lambda$, then $u_j(CER_j(e)) = u_j(\lambda) < u_j(\sigma_j(1)) = 1 = u_i(CER_i(e))$.

Case 2: $W > \sum_{i \in I} \sigma_i(1)$

In this case, $CER_i(e) < CER_j(e)$ if and only if $CER_j(e) = \sigma_j(1)$ and $CER_i(e) = \min\{\sigma_i(1), \lambda\} < \sigma_j(1)$, where λ is such that $\sum_{l \in I} \min\{\sigma_l(1), \lambda\} = W$. If $CER_i(e) = \sigma_i(1)$, then $u_j(CER_j(e)) = 1 = u_i(CER_i(e))$. If $CER_i(e) = \lambda$, then $\sigma_j(1) > CER_i(e) \geq \sigma_i(1)$ and therefore $u_j(CER_j(e)) = u_j(\sigma_j(1)) = 1 = u_i(\sigma_i(1)) \leq u_i(CER_i(e))$.

- CER satisfies CY

Let $e = (I, u, W) \in \mathcal{E}$. Let $I' \subset I$ with $|I'| = 2$ and $e' = (I', u', W')$, where $u' = (u_i)_{i \in I'}$ and $W' = \sum_{i \in I'} CER_i(e)$. Two cases are distinguished.

Case 1: $W \leq \sum_{i \in I} \sigma_i(1)$.

In this case, $CER_i(e) = \min\{\sigma_i(1), \lambda\}$ for all $i \in I$, where λ is such that $\sum_{i \in I} \min\{\sigma_i(1), \lambda\} = W$. Thus, $CER_i(e) \leq \sigma_i(1)$ for all $i \in I$ and therefore $W' \leq \sum_{i \in I'} \sigma_i(1)$. Consequently, $CER_i(e') = \min\{\sigma_i(1), \lambda'\}$ for all $i \in I'$, where λ' is such that $\sum_{l \in I'} \max\{\sigma_l(1), \lambda'\} = W'$. If $\lambda = \lambda'$, there is nothing to prove. Assume then $\lambda < \lambda'$. The proof for $\lambda > \lambda'$ is analogous. In this case, $CER_i(e) \leq CER_i(e')$ for all $i \in I'$. If there would exist some $i_0 \in I'$ such that $CER_{i_0}(e) < CER_{i_0}(e')$, then $W' = \sum_{i \in I'} CER_i(e) < \sum_{i \in I'} CER_i(e') = W'$, which is a contradiction. Thus, $CER_i(e) = CER_i(e')$ for all $i \in I'$.

Case 2: $W > \sum_{i \in I} \sigma_i(1)$.

In this case, $CER_i(e) = \max\{\sigma_i(1), \lambda\}$ for all $i \in I$, where λ is such that $\sum_{i \in I} \max\{\sigma_i(1), \lambda\} = W$. Thus, $CER_i(e) \geq \sigma_i(1)$ for all $i \in I$ and therefore $W' \geq \sum_{i \in I'} \sigma_i(1)$. Consequently, $CER_i(e') = \max\{\sigma_i(1), \lambda'\}$ for all $i \in I'$, where λ' is such that $\sum_{i \in I'} \max\{\sigma_i(1), \lambda'\} = W'$. The rest of the proof is similar to that of Case 1.

- CER does not satisfy RM

Let $u_1(x) = x$ and $u_2(x) = x/10$, for all $x \in \mathbb{R}_+$. Then, $\sigma_1(x) = x$ and $\sigma_2(x) = 10 \cdot x$, for all $x \in \mathbb{R}_+$. Consider $e = (\{1,2\}, \{u_1, u_2\}, 5)$ and $e' = (\{1,2\}, \{u_1, u_2\}, 6)$. Then, $CER(e) = (1,4)$ and $CER(e) = (1,5)$. ■

6.3.3 Strong Priority

Finally, let P be the rule that allocates the wealth of an economy proportionally to agents' claims in the economy. Formally, for all $e = (I, u, W) \in \mathcal{E}$,

$$P(e) = \lambda(e) \cdot (\sigma_i(1))_{i \in I},$$

where $\lambda(e) \in \mathbb{R}_+$ is such that $P(e) \in S(e)$, i.e., $\lambda(e) = \frac{W}{\sum_{i \in I} \sigma_i(1)}$.

Proposition 6 *P satisfies RM and CY but it fails to satisfy TP.*

Proof.

It is straightforward to see that P satisfies RM. We show now that P satisfies CY. Let $e = (I, u, W) \in \mathcal{E}$. Let $I' \subset I$ with $|I'| = 2$ and $e' = (I', u', W')$, where $u' = (u_i)_{i \in I'}$ and $W' = \sum_{i \in I'} P_i(e)$. For all $i \in I'$,

$$P_i(e) = \frac{W}{\sum_{i \in I} \sigma_i(1)} \cdot \sigma_i(1) \text{ and } P_i(e') = \frac{W'}{\sum_{i \in I'} \sigma_i(1)} \cdot \sigma_i(1).$$

Since $W' = \sum_{i \in I'} P_i(e) = \frac{W}{\sum_{i \in I} \sigma_i(1)} \cdot \sum_{i \in I'} \sigma_i(1)$, it follows that $P_i(e') = \frac{W}{\sum_{i \in I} \sigma_i(1)} \cdot \sigma_i(1) = P_i(e)$.

Finally, P does not satisfy TP. To show this, for all $x \in \mathbb{R}_+$, let $u_1(x) = x/2$ and $u_2(x) = x^2$. Consider $e = (\{1,2\}, \{u_1, u_2\}, 2)$. Then, $P(e) = (4/3, 2/3)$ and $u_1(4/3) = 2/3 > 4/9 = u_2(2/3)$. ■

6.4 Solidarity and Consistency

Finally, we show the logical relation among solidarity axioms and consistency mentioned in Section 4.2. First, consider the following definition:

Consistency *(CY^*). Let $e = (I, u, W) \in \mathcal{E}$. Let $I' \subset I$ and $e' = (I', u', W')$, where $u' = (u_i)_{i \in I'}$ and $W' = \sum_{i \in I'} F_i(e)$. Then $F_i(e) = F_i(e')$, for all $i \in I'$.*[26]

Proposition 7 *A rule satisfies AG if and only if it satisfies RM and CY^*.*

Proof.

Let F be an allocation rule satisfying AG. It is straightforward to show that F satisfies RM. Let us see that it also satisfies CY^*. Let $e = (I, u, W) \in \mathcal{E}$ be given and fix $I' \subset I$. Denote $w = (F_i(e))_{i \in I'}$ and consider the new economy $e' = (I', u', W')$, where $u' = (u_i)_{i \in I'}$ and $W' = \sum_{i \in I'} F_i(e)$. By AG we either have $F(e') = w$, $F(e') < w$ or $F(e') > w$. Now, since F is an allocation rule then it allocates the wealth of each economy completely, i.e., $\sum_{i \in I'} F_i(e') = W' = \sum_{i \in I'} F_i(e)$. Thus, $F(e') = w$, which shows that F satisfies CY^*.

Let now F be an allocation rule satisfying RM and CY^*. Let $e = (I, u, W)$ and $e' = (I', u', W') \in \mathcal{E}$, such that $I' \subseteq I$. Consider the auxiliary economy $\widehat{e} = (I', u', W) \in \mathcal{E}$. Note that $W = \sum_{i \in I} F_i(e)$. Thus, by CY^*, $F(\widehat{e}) = F_{I'}(e)$. Finally, depending on the relationship between W and W', and thanks to RM, one of the following three possibilities happens: $F(\widehat{e}) = F(e')$, $F(\widehat{e}) > F(e')$ or $F(\widehat{e}) < F(e')$. This concludes the proof. ■

[26] Note that if $|I'| = 2$ this coincides with the axiom CY introduced in Section 4.

References

[1] Barry, B. (1989), **Theories of Justice**, Vol. 1. Berkeley. University of California Press.

[2] Barry, B. (1995), **Justice as impartiality**, Oxford: Clarendon Press

[3] Binmore, K. (1994), **Playing fair**, MIT Press. Cambridge, MA.

[4] Cohen, G.A. (1992), "Incentives, inequality, and community," in Grethe Peterson, ed., The Tanner Lectures on Human Values, Salt Lake City: University of Utah Press

[5] Chun, Y. (1986), The solidarity axiom for quasi-linear social choice problems, **Social Choice and Welfare** 3, 297-310.

[6] Chun, Y. (1999), Equivalence of axioms for bankruptcy problems, **International Journal of Game Theory** 28, 511-520.

[7] Dworkin, R. (1981a), What is equality? Part 1: Equality of welfare. **Philosophy & Public Affairs** 10, 185-246.

[8] Dworkin, R. (1981b), What is equality? Part 2: Equality of resources. **Philosophy & Public Affairs** 10, 283-345.

[9] Ehlers, L. and Klaus, B. (2001), Solidarity and probabilistic target rules, **Journal of Public Economic Theory** 3, 167-184.

[10] Ehlers, L. and Klaus, B. (2004), Resource-Monotonicity for House Allocation Problems, **International Journal of Game Theory**, forthcoming.

[11] Fleurbaey, M. and Maniquet F. (1999), Cooperative production with unequal skills: The solidarity approach to compensation, **Social Choice and Welfare** 16: 569-583

[12] Fleurbaey, M. and Maniquet, F. (2004) "Compensation and responsibility," in K. Arrow, K. Suzumura, and A. Sen (eds.), Handbook of Social Choice and Welfare (in press).

[13] Harsanyi, J. (1953), Cardinal utility in welfare economics and in the theory of risk-taking, **Journal of Political Economy** 61, 434-435.

[14] Klaus, B. and Miyagawa, E. (2001), Strategy-proofness, solidarity, and consistency for multiple assignment problems, **International Journal of Game Theory** 30, 421–435.

[15] Keiding, H. and Moulin H. (1991), The solidarity axiom in parametric surplus-sharing problems, **Journal of Mathematical Economics** 20, 249-270.

[16] Kolm, S.-C. (2004), Reciprocity: its scope, rationale, and consequences. **Mimeo**, available at http://aran.univ-pau.fr/ee/

[17] Moulin, H. (2003), **Fair division and collective welfare**, MIT Press. Cambridge, MA.

[18] Parfit, D. (1997), **Equality and Priority**. Ratio: New Series 10, 202-221. Blackwell Publishers Ltd. Oxford.

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

[20] Roemer, J.E. (1986), Equality of resources implies equality of welfare. **Quarterly Journal of Economics** 101, 751-784.

[21] Roemer, J.E. (1996), **Theories of Distributive Justice**, Harvard University Press, Cambridge MA.

[22] Roemer, J.E. (2002), "Egalitarianism against the veil of ignorance," **Journal of Philosophy** 99, 167-184.

[23] Roemer, J.E. (2004), "Eclectic distributional ethics," **Philosophy, Politics, and Economics** (in press).

[24] Scanlon, T. (1998), **What we owe to each other**, Harvard University Press. Cambridge, MA.

[25] Sen, A. (1973), **On Economic Inequality**, Clarendon Press, Oxford.

[26] Sen, A. (1980), "Equality of what?" in S. McMurrin, ed. The Tanner Lectures on Human Values. Vol 1. Salt Lake City: University of Utah Press.

[27] Sen, A. (1992), **Inequality Reexamined**, Harvard University Press. Cambridge, MA.

[28] Sprumont, Y. (1996), Axiomatizing Ordinal Welfare Egalitarianism When Preferences May Vary, **Journal of Economic Theory** 68, 77-110.

[29] Temkin, L.S. (1993), **Inequality**, New York: Oxford University Press.

[30] Thomson, W. (1983), Equity in Exchange Economies. **Journal of Economic Theory** 29, 217-244.

[31] Thomson, W. (1995), Monotonic allocation rules, University of Rochester Discussion Paper.

[32] Thomson, W. (2004), Consistent Allocation Rules, Book Manuscript.

[33] Thomson, W. and T. Lensberg (1989), **Axiomatic Theory of Bargaining with a Variable Number of Agents.** Cambridge University Press, New York.

[34] Young, P., (1994). **Equity**, Princeton University Press, Princeton.

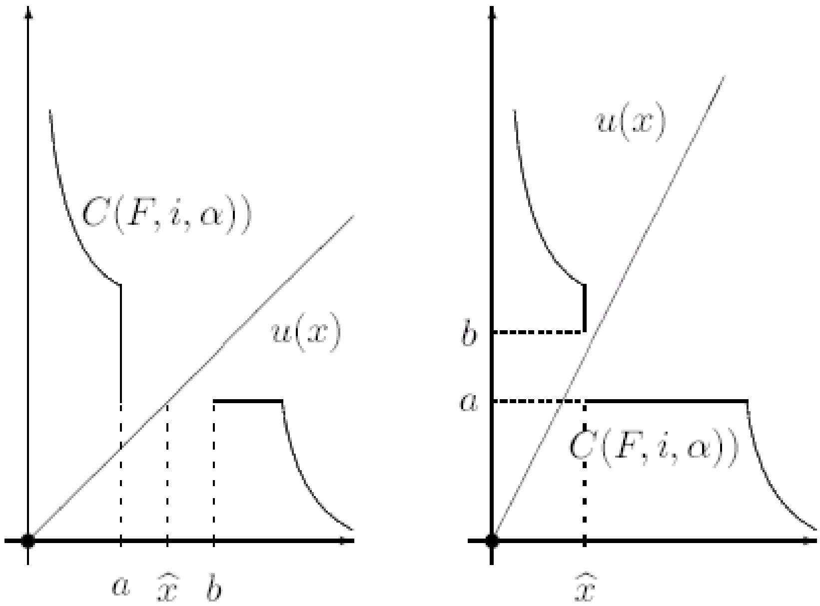

Figure 1: Proof of Lemma 5.

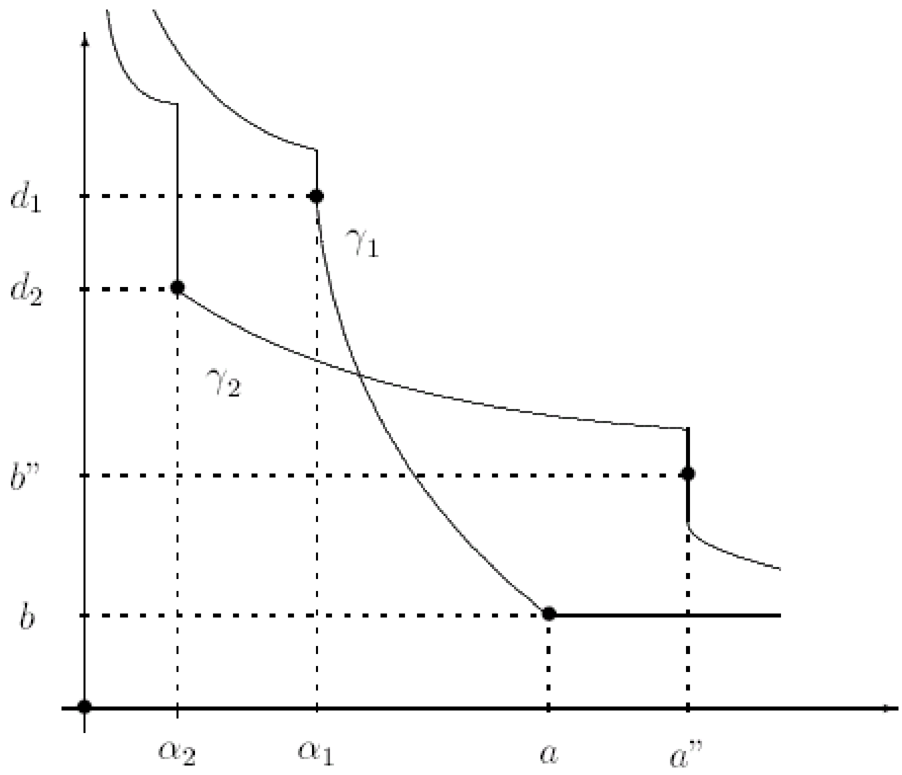

Figure 2: Proof of Lemma 7