

**Distributive justice and benevolence:  
The welfare state as practical, distributive, liberal social contract<sup>1</sup>**

Jean Mercier Ythier  
Sorbonne Universités  
Université Panthéon-Assas  
CRED and IDEP  
(January 2012, revised February 2013)

**Abstract:** Benevolence is the humanist virtue par excellence. Influential veil-of-ignorance theories of justice, issuing from the main streams of the humanist ethics of the Enlightenment, exclude benevolence from the short list of individual motives from which they derive their norms of justice. These norms of justice are built from hypothetical preferences of individuals, while the institutions of distributive justice in real-life democracies express, rely on, or are derived from individuals' actual preferences. Democratically implementable norms of justice should be derived from individuals' actual preferences and from suitable norms of deliberation and communication in and through collective action. They are norms of benevolent justice, in the sense that they rely significantly on the altruistic motives of individuals. Contemporary welfare states are interpreted as practical achievements of such a norm of benevolent justice within the distributive liberal social contract.

*Keywords:* benevolence, distributive justice, democracy, welfare state

**Justice distributive et bienveillance:  
L'Etat-Providence en tant que réalisation pratique du contrat social libéral**

**Résumé:** La bienveillance est la vertu humaniste par excellence. D'importantes théories de la justice "sous le voile d'ignorance", issues des principaux courants de l'éthique humaniste des Lumières, excluent la bienveillance de la liste des raisons et motifs individuels sur lesquels elles fondent leurs normes de justice. Ces normes de justice sont construites sur la base de préférences hypothétiques des individus, alors que les institutions de justice distributive des démocraties reposent, dans la réalité pratique de ces systèmes politiques, sur les préférences effectives des individus. Nous défendons l'idée que les normes de justice, pour être applicables en démocratie, doivent reposer sur les préférences effectives des citoyens, s'exprimant dans le cadre d'un fonctionnement politique obéissant à des normes adéquates de délibération et de communication dans l'action collective et par celle-ci. Ce sont des normes de justice bienveillante, au sens où elles reposent de manière significative sur les raisons et motifs altruistes des individus. Les Etats-Providence contemporains sont interprétés ici comme le résultat de la mise en œuvre pratique d'une norme de justice bienveillante de ce type, celle du contrat social libéral.

*Mots-clés:* bienveillance, justice distributive, démocratie, Etat-Providence

JEL classification codes: H10, I30, Z13

---

<sup>1</sup> I warmly thank Christina Fong, Stefano Zamagni and two anonymous referees for helpful comments. I also thank participants in the ERMES-LEM seminar at Université Panthéon-Assas, and in the 2013 meeting of the European Public Choice Society at Zürich, for useful discussions at two stages of my elaboration of the paper. The detailed report of one referee proved particularly stimulating. I wrote section 5 to address some of the main issues raised by that reader. Any errors, of course, remain my responsibility.

## 1-Introduction

Aristotle's *Nicomachean Ethics* devotes its whole fifth book to the virtue of justice, which is also one of the four cardinal virtues of Thomas Aquinas's *Summa Theologica*. Adam Smith extends the tradition further, by including justice as one of his three main virtues in the *Theory of Moral Sentiments*: "The man who acts according to the rules of perfect prudence, of strict justice, and of proper benevolence, may be said to be perfectly virtuous" (Part VI, Section III, p. 349).<sup>2</sup> Benevolence doesn't appear among the virtues enumerated in the *Nicomachean Ethics*; its closest analogues in that work are liberality and magnificence, which Aristotle considers in the first and second chapters of Book IV, respectively. Nor does benevolence belong to Aquinas's list of the cardinal and theological virtues, its closest analogue being the theological virtue of charity. The differences between Aristotle's liberality, charity as a theological virtue, and benevolence point to deeper differences between these three thinkers' ethical systems. Placing benevolence on a short list of basic virtues means that it is *fundamentally* good, for humane reasons, to wish the good for other human beings. Benevolence is the humanist virtue par excellence.<sup>3</sup>

Among the theories of justice that exerted the widest influence in the field of economics in the second half of the twentieth century, Harsanyi's utilitarianism (1953, 1955, 1977) and Rawls's *Theory of Justice* (1971) differ from the virtue-ethical tradition(s) on at least two fundamental grounds. First, they tend to develop methodologically as a separate body of normative ethics, concentrating specifically on the ethical norms of political and social organization and action. They constitute an *analytic* development within the field of normative ethics. Second, on substantive grounds, their theories extend the two main contributions of the Enlightenment to normative ethics, namely, utilitarian and Kantian ethics, respectively.

Harsanyi and Rawls arrive at their quite different conceptions of justice from the application of (some principles of) reason under a hypothetical *veil of ignorance* relative to individuals' characteristics and positions in society. Alternative modern conceptions construe justice, by contrast, as the balanced outcome (if one is possible at all) of reasoned deliberation and communication by actual persons in society. The former approach excludes benevolent feelings, and thus the corresponding virtue of benevolence, from the derivation of justice, whereas the latter permits a role for benevolence in the definition of justice.

We will analyze below the consequences of a hypothetical veil of ignorance for the conception of justice and its relation to benevolence (sections 2 and 3). We will then move on to consider theories of benevolent justice that derive their conceptions of justice from the actual preferences of individuals, including notably their actual benevolent motives, and from suitable norms of deliberation and communication in and through collective action (section 4). Finally, we will develop an interpretation of contemporary welfare states as practical achievements of the norm of benevolent justice within the distributive liberal social contract (section 5).

---

<sup>2</sup> The common elements of the lists of virtues of Aristotle, of the doctors of the Church, and of Adam Smith are justice and prudence. See Mercier Ythier (2013) for an account of the influence of virtue ethics in economic science.

<sup>3</sup> Smith's ethics appears quite modern in this respect, although references to God remain quite apparent in the *Theory of Moral Sentiments* (see for example Part VI, Section II, Chap. III). Alasdair MacIntyre (2007, p. 234) characterizes Adam Smith as "a deist rather than a Christian".

## 2-Justice from impartial reason

As already noted, both Harsanyi (1953, 1955) and Rawls (1971) used the veil of ignorance as an analytic device, in order to define a hypothetical position of *impartiality*, from which they could proceed to the rational derivation of their conceptions of justice.<sup>4</sup> The two authors develop two competing interpretations of impartiality (and, consequently, two competing conceptions of justice): Harsanyi interprets impartiality (and justice) as *impersonality*, while Rawls characterizes it as *fairness*.

The situations of justice considered in these theories are ones that require judging between competing claims of individuals concerning their respective life prospects. The veil of ignorance creates impartiality by enabling individuals to ignore all the circumstances that may condition their own life prospects, such as inherited wealth or health, individual talents, motives or preferences, and one's natural or social environment.

Harsanyi (1953, 1977) further interprets this position of ignorance in terms of *objective probabilities and utilities*. Ignorance is converted into a uniform (objective)<sup>5</sup> probability distribution over events (which consist of "being such person endowed with such set of individual circumstances"), by means of Laplace's principle of insufficient reason. The preferences of the impartial observer are defined over lotteries, the consequences of which consist of "being such person in such state of the world".<sup>6</sup> These preferences verify the axioms of the theory of expected utility of von Neumann and Morgenstern: that is, the preferences above not only are rational, but also verify suitable continuity properties and the axiom of independence. The impartiality of the observer translates into the following additional assumption: a state of the world (say, state  $x$ ) is preferred to another (say, state  $y$ ) if and only if the lottery giving all individuals the same chance of experiencing state  $x$  is preferred to the lottery giving all individuals the same chance of experiencing state  $y$ . Finally, and most importantly perhaps, it is assumed that the preferences of the impartial observer relative to consequences  $(i, x)$  (being person  $i$  in state of the world  $x$ ) coincide with individuals' preferences relative to states of the world: that is, there exists a utility representation  $u$  of the impartial observer's preferences such that  $u_i$  defined by  $u_i(x) = u(i, x)$  for all  $(i, x)$  is a utility representation of  $i$ 's personal preferences relative to  $x$

---

<sup>4</sup> The first analytic use of the veil of ignorance is usually attributed to Vickrey (1945).

<sup>5</sup> A probability distribution is *objective* if it is endorsed by all individuals. The theory of expected utility of von Neumann and Morgenstern (1944) supposes objective probabilities, while the variant of this theory developed by Savage (1954) allows for *subjective* probabilities, that is, for possible differences in individuals' opinions relative to the probabilities of events. Savage's axioms of rational decision under uncertainty imply that the preference relation of an individual embodies his or her individual evaluation of the probabilities of events: that is, one's expected utility is computed from one's (subjective) probability distribution.

<sup>6</sup> A lottery is defined formally as a probability distribution over a set of definite consequences (the lottery "prizes", which can be lotteries themselves in the case of compound lotteries). The account below follows those of Weymark (1991) and Fleurbaey (1996, section 4.4.2). An alternative recent account is provided by Grant et al. (2010), who give axiomatic characterizations of: (i) *generalized utilitarianism*, that is, to use our notation, of preferences of the impartial observer represented by a utility function of the type  $(1/n)\sum_{i=1}^n \varphi_i(u_i(x))$ ; (ii) *generalized utilitarianism with a common function  $\varphi$* , yielding a utility function of the impartial observer of the type  $(1/n)\sum_{i=1}^n \varphi(u_i(x))$ ; (iii) and *utilitarianism*, where the preferences of the impartial observer admit a representation of the type  $(1/n)\sum_{i=1}^n u_i(x)$ . They show, in particular, that utilitarianism obtains from generalized utilitarianism if and only if the impartial observer is indifferent between life chances (lotteries relative to social outcomes  $x$ ) and accidents of birth (lotteries relative to personal identities  $i$ ).

(this is Weymark's identity axiom, which is analogous to but stronger than Harsanyi's acceptance principle<sup>7</sup>). These assumptions together imply that the preferences of the impartial observer admit a utility representation of the type  $(1/n)\sum_{i=1}^n \varphi(u_i(x))$ , where  $\varphi: \mathbb{R} \rightarrow \mathbb{R}$  is strictly increasing.

Rawls interprets the position of ignorance in Kantian objective terms. He rejects the principle of insufficient reason as lacking objective grounds in "known features of one's society" (1971, III, 28, p. 168), and views the position of ignorance as non-probabilizable altogether<sup>8</sup>. He justifies, likewise, by a suitable objective reason his use of primary goods (rights and liberties, opportunities and powers, income and wealth notably) as the basis of the (objective) appraisal of the good of individuals: "[...] primary goods [...] are things which it is supposed a rational man wants whatever else he wants [...]. With more of these goods rational men can generally be assured of greater success in [...] advancing their ends, whatever these ends may be" (1971, II, 15, p. 92).<sup>9</sup> The position of ignorance takes place in a wider notion of original

---

<sup>7</sup> Grant and al. (2010, section 5, Proposition 6) show that Weymark's identity axiom is equivalent to the combination of Harsanyi's acceptance principle with the impartial observer's indifference between individuals facing similar risks. Weymark's axiom and the still stronger axiom of indifference between life chances and accidents of birth (which yields utilitarianism, as discussed in the previous footnote) both imply a high degree of interpersonal comparability of welfare. With such assumptions we are not in a subjective utility framework anymore. The impersonal character of Harsanyi's impartiality finds its exact expression in the supposed ability of any individual to assess the welfare of any other, either ordinally (Weymark) or cardinally (utilitarianism), from the same common preference relation over "being such person in such state of the world". This impersonal common preference relation (if any) confers a character of objectivity to interpersonal comparisons of welfare, hence to the evaluation of individual welfare itself. Note that the notion of impersonal objectivity so implied is conceptually distinct from the measurability of individual welfare. It simply refers to the existence of a common standpoint (if any) for making such statements as "being individual  $i$  in state of the world  $x$  is better than being individual  $j$  in state of the world  $y$ " or "being individual  $i$  in state of the world  $x$  is better than being (the same) individual  $i$  in state of the world  $y$ ". From the common impersonal standpoint, all individuals would agree on both statements. Moreover, individual  $i$  would agree to the second statement considered from his or her *personal* standpoint, and would understand it as equivalent to saying "I prefer state of the world  $x$  to state of the world  $y$ ". Rawls seems to reject the impersonal standpoint as meaningless (1971, III, 28), and clearly rejects impersonality as an appropriate definition of impartiality (1971, III, 30). A penetrating philosophical (metaphysical) exposition and defense of impersonality and related notion of objectivity may be found in Nagel (1970, part III, chapters 10-12).

<sup>8</sup> This is implied by Rawls's defense of the difference principle as opposed to the average utility principle (1971, III, 28). An intermediate position between Rawls (no probabilities at all) and Harsanyi (objective uniform probabilities) would use Savage's expected utility theory (subjective probabilities: see footnote 5 above). Savage's theory would presumably imply the existence of as many impartial observers as there are distinct individuals (supposing that individual probability distributions are pairwise distinct), even if one assumes the existence of an impersonal standpoint as discussed in footnote 7.

<sup>9</sup> Note the shift in the notions of objectivity mobilized to criticize the principle of insufficient reason, on one hand, and to substitute primary goods for individual welfare on the other hand. The first notion of objectivity refers to the accurate representation of social reality, while the second one refers to unanimous reasoning (i.e. unanimous *minds*). Our interpretation of Harsanyi above exactly reverses the use of the same notions, by letting Laplace's principle be unanimously accepted, and by supposing that the impersonal standpoint conveys the true representation of the real preferences of individuals (i.e. that it accurately represents their actual minds). Note, also, that Rawls's conception of individual preferences relative to primary goods resorts to both of the conceptions of objectivity outlined above, namely, objectivity as convention (i.e. objectivity from unanimous agreement), and objectivity as truthful representation of some suitable part of the reality of individual psychologies (the part that, according to Rawls, expresses the nature of human beings as free and equal rational beings). The natural sciences, as bodies of knowledge, understand objectivity only in the second sense (truthful representation). But they do rely implicitly on the first notion of objectivity as human practices, if only through the conventional agreement that truthful representation is the appropriate aim of their epistemic activity. The social sciences, whose object of study is human practices, necessarily make a *joint* use of the two notions of objectivity, with deep consequences for the epistemic notion of "scientific truth" that they mobilize.

position, which consists of a hypothetical position of liberty and equality of all individuals in the “circumstances of justice”, that is, in the hypothetical circumstances of their participation in the elaboration of the unanimous social contract that is supposed to produce the norm of justice. The original position is construed, more precisely, as the set of conditions that allows for the full expression of the nature of human persons as “free and equal rational beings with a liberty to choose” (1971, IV, 40, p. 256)<sup>10</sup>. Human nature, so interpreted in Kantian terms, requires not only the definition of the original position as just suggested, but also the principles of justice that obtain from it, by rational deliberation, at the “reflective equilibrium” (1971, I, 4, p. 20) of the social contract.

Rawls’s original position differs from Harsanyi’s impartial impersonal standpoint in two main respects: (i) on *informational* grounds, in that it views primary goods as the appropriate basis for the appraisal of situations of justice, and (ii) on *motivational* grounds, as it assumes the mutual disinterest of individuals. This second feature means that each individual grounds his or her appraisal of the circumstances of justice under the veil of ignorance on his or her own endowment of primary goods in all conceivable individual life prospects. This motivational feature of the original position maximizes the possibilities of expression, by every individual, of his or her universal (“categorical”) reasons (that is, his or her preferences relative to primary goods), in the sense that the expression of mutually disinterested individual preferences is not limited by the consideration of the claims of others. The informational and motivational assumptions embodied in Rawls’s original position suppose, in other words, the *full* (equal) *liberty* of individuals to express their *universal* individual ends, whereas the informational and motivational assumptions embodied in Harsanyi’s impartial impersonal standpoint presuppose people’s *impartial imaginative empathy* relative to their *contingent* individual ends.

Summing up, for Rawls, the *informational basis* of justice consists of a substantive list of primary goods supporting human nature as free and equal rational beings; for Harsanyi it lies in common preferences relative to “being such person in such state of the world”. The *motivational basis* of justice, for Rawls, consists of individuals’ (identical) preferences relative to their own endowment of primary goods in all conceivable real circumstances, whereas for Harsanyi it entails impartial impersonal identification with all personal individual preferences. Moreover, the norm of justice obtains from the corresponding informational and motivational basis (that is, from the *original position* on the one hand, and from the *impartial impersonal standpoint* on the other hand) through different methods of rational deliberation: unanimous agreement at reflective equilibrium, in Rawls’s conception; unanimous agreement about probabilities and about the axioms of decision under uncertainty of von Neumann and Morgenstern, in Harsanyi’s conception.

Benevolent sentiments, and hence benevolence as a virtue, are excluded by definition from both the impartial impersonal standpoint and the original position. In Rawls’s case, mutual disinterestedness eliminates benevolence from the motivational basis of the original position. The ties of benevolent sentiments put limits on individual autonomy as construed by Rawls, particularly if these ties proceed from the so-called sentiments of natural affection—that is, if they are the affective expression of membership in some specific community (such as a family, or a religious or ethnic group) that is a subset of the political unit constituted by the social contract. They belong, in other words, to the contingent (“hypothetical”, in Kantian

---

<sup>10</sup> Also: “The description of the original position is an attempt to interpret this conception”, that is, Kant’s conception of moral principles as “legislation for a kingdom of ends [...] to be agreed to under conditions that characterize men as free and equal rational beings” (1971, IV, 40, p. 252).

vocabulary) circumstances of individual life that disappear altogether under the Rawlsian veil of ignorance, as irrelevant (usually partial) information and reason. Benevolent sentiments do belong to the informational basis of Harsanyi's impartial impersonal standpoint, but they are absent, as sentiments, from its motivational basis. The imaginative empathy of Harsanyi's impartial observer is axiomatically rational, in the sense that it is deduced from the application of axioms of rational decision making relative to generic lotteries, the definition of which is independent of the nature of lottery "consequences" (the same axioms apply whether the "prize" is one's personal identity or one's stakes in a poker game). Harsanyi's empathy is *not* the (both imaginative and sentimental) sympathy of Adam Smith, of classical utilitarianism<sup>11</sup>, or, more generally, of the ethical tradition of British sentimentalism. Sentiments belong to the personal self of individuals in Harsanyi's construct. They disappear, accordingly, under the Harsanyian veil of ignorance, as irrelevant (that is, personal) motives, sources of partiality of individual judgments.

### 3-Justice as equality of resources

Dworkin's notion of distributive justice as equality of resources (1981a, 1981b) occupies a singular position in the present account. It is posited on the knife edge between the conceptions of justice that we have discussed in section 2, derived from hypothetical individual preferences that are both identical and impartial, and those to be examined in section 4, derived from the communication between individuals whose actual (as opposed to hypothetical) preferences entail some form of partiality, stemming from individuals' circumstances or motives. The main characteristic features of Dworkin's construct, for the purpose of this classification, are (i) that individual preferences are constructed under a hypothetical veil of ignorance, but one so "thin" that they remain personal and partial, and (ii) that the derivation of the norm of justice from these conflicting hypothetical preferences relies on an ideal communication process, which forms therefore an integral part of this definition of justice.

Central to Dworkin's construct is the distinction between individuals' resources and preferences: "The distinction required by equality of resources is the distinction between those beliefs and attitudes that define what a successful life would be like, which the ideal assigns to the person, and those features of body or mind or personality that provide means or impediments to that success, which the ideal assigns to the person's circumstances" (1981b, p. 223).

The preferences that are considered here are the *personal* preferences of the individual, which Dworkin distinguishes from two other categories: one's *impersonal* preferences, such as one's beliefs regarding impersonal concepts like scientific progress or the protection of biodiversity, and one's *political* preferences—e.g. one's views relative to political action or institutions, or with regard to interpersonal relations, benevolent or otherwise (1981a). Dworkin uses the concept of personal preferences of an individual to mean *mutually disinterested* preferences

---

<sup>11</sup> We adopt the definition of classical utilitarianism given by Rawls (1971, I, 5), including the references in his footnote 9, pp. 22-23, with two exceptions. First, Adam Smith's *Theory of Moral Sentiments* does *not* endorse the principle of utility, understood as the grounding of ethical evaluation on the total sum of individuals' welfare: see Mercier Ythier (2013); second, Harsanyi substitutes rational empathy for (sentimental) sympathy. For conceptual clarity we propose to reserve, in this paper, the phrase "classical utilitarianism" for the main traditional stream of utilitarian ethics, which both endorses the principle of utility and interprets the principle as expressing sentiments of (universal) benevolence. Classical utilitarianism, so characterized, includes (but is not limited to) Hutcheson, Hume, Beccaria, John Stuart Mill and Edgeworth, and its substance is summarized in Sidgwick's *Methods of Ethics* (1907).

relative to the *personal* ends of that individual. They define the good life from the individual's isolated (personal, unsympathetic) standpoint. An individual's ("extended") resources are the means provided by the person's circumstances, for achieving the aims implied by her or his personal preferences so conceived. They include not only his or her material resources, but also his or her other physical and mental powers such as determined by his or her talents, skills or disabilities.

Dworkin's norm of justice, namely, equality of resources, is defined from a hypothetical initial position characterized as follows. First, individuals know their personal preferences (in Dworkin's sense) but they ignore their extended resources: that is, individuals' resources, and notably their talents and handicaps, are hidden behind a veil of ignorance, which is nevertheless too "thin" to hide their personal preferences. Second, their wealth in the initial position reduces to equal amounts of money "numeraire" (i.e. money), which can be used for exchanges on a complete system of competitive markets, including insurance markets for handicaps and talents. Third, individuals' preferences under uncertainty, which determine their insurance choices under the veil of ignorance, are deduced from their personal preferences by means of Laplace's principle of insufficient reason and the axioms of decision making under uncertainty of von Neumann and Morgenstern.<sup>12</sup> Equality of resources consists of any Walrasian equilibrium of this hypothetical economy.<sup>13</sup>

Summing up, the *informational* basis of Dworkin's norm of justice consists of individuals' personal (mutually disinterested) preferences. Its *motivational* basis consists of the personal insurance motives embodied in the expected personal utilities of individuals computed as indicated above. These informational and motivational features are complemented by, and articulated in accordance with, the procedural and positional features of the Walrasian auction with equal money endowments of individuals.

#### 4-Benevolent justice

The two conceptions of justice in section 2 confront the same basic difficulty, namely, a tension between the global character of the ends of justice, as pertaining to the whole life prospects of all individuals in society, and its grounding on the sole impartial reason of individuals. In actual democratic societies, the institutions of distributive justice, and notably

---

<sup>12</sup> More specifically, contingent states of the hypothetical initial position consist of definite assignments of extended resources to individuals. Let us designate them by an index  $e$  running in  $\{1, \dots, m\}$ ,  $m \geq 2$ . Let  $u_i(x_e^i)$  denote individual  $i$ 's personal utility from her contingent consumption bundle  $x_e^i$  in state  $e$ ,  $r_e^i$  denote individual  $i$ 's contingent money wealth in state  $e$ , and  $p_e$  denote the system of contingent market prices of state  $e$ . Individual  $i$ 's indirect utility in state  $e$  reads  $v_i(p_e, r_e^i) = \max\{u_i(x_e^i) : p_e x_e^i \leq r_e^i\}$ . The functioning of Dworkin's hypothetical economy may be modelled, in a standard way, as a two-period competitive economy where: in the first period, individuals ignore their extended resources and spend their equal endowments of money on insurance services so as to maximize their expected utilities  $\sum_{e=1}^m \frac{1}{m} v_i(p_e, r_e^i)$  with respect to random income distribution  $(r_1^i, \dots, r_m^i)$ ; in the second period, individuals know the contingent state  $e$ , and spend their contingent income  $r_e^i$  on consumption commodities, purchased at the contingent price system  $p_e$ , so as to maximize their personal utility  $u_i(x_e^i)$ .

<sup>13</sup> We concentrate, here, on the second and richer of Dworkin's two definitions of equality of resources, namely, his scheme of *fundamental insurance* (1981b, III-VI). The first definition (1981b, I) is simply the Walrasian equilibrium in certainty with equal money wealth of individuals (which implies, notably, that no individual envies the allocation of others from the standpoint of his or her personal preferences).

the institutions of the welfare state, must rely on individuals' *actual* preferences, political or otherwise. This observation leaves us with a simple alternative for the theoretical derivation of democratically implementable norms of justice. One can either maintain the global ends of justice, and seek to enrich its motivational foundations, notably by including benevolence in the latter; or, alternatively, one can maintain impartial reason (however defined) as the motivational foundation of justice, and narrow the scope of the ends of justice (for instance, by limiting it to the laws of a society, defined in a narrow or exclusive way<sup>14</sup>). The sampling of theories of justice reviewed below explores the first of these two options.

Dworkin's norm of distributive justice faces essentially the same type of difficulty as the constructs of Rawls and Harsanyi. The norm is derived from two sources: (i) competing insurance motives of mutually disinterested individuals confronting a fundamental (hypothetical) uncertainty relative to their extended resources, and (ii) a Walrasian auction with equal money endowments of individuals, as a fair initial position and an ideal communication process. The question may be raised, then, following Dworkin's own distinctions relative to individuals' actual preferences, whether their political preferences (using that term in his sense) would support the norm of justice so derived from their personal, mutually disinterested preferences. My guess is that they presumably would not (and they actually *do not* in existing political systems) for distributive justice *as a whole*, but that they presumably would support, at least in principle, the component of distributive justice that concerns the compensation for major non-insurable accidents such as congenial birth disabilities for example. To put it in less abstract terms, there is, in my opinion, a clear case for political support of the *right* to compensation for non-insurable accidents at birth or in life, based partly (if not exclusively) on widespread fundamental insurance motives. But this case does not extend to the whole of distributive justice, as construed in the all-inclusive terms of contemporary theories of justice.<sup>15</sup>

Justice as impartial reason is the expression of a suitably defined (unanimous) individual reason. The unanimity of impartial motives rules out communication and collective deliberation as independent sources of principles of justice: under the (thick) veil of ignorance, individual and collective deliberation, communication, and individual and collective choice are but the several aspects of one and the same construction, namely, the identical views of individuals relative to justice. Lifting up the veil of ignorance, either partly (as in Dworkin's construct) or completely (as in the theories of justice considered below) drives a wedge between individual reason and social choice or communication. It thus causes the conditions of social communication and associated practices of collective deliberation and

---

<sup>14</sup> This is the perspective adopted, for instance, by many contributions to the domain of "law and economics". Considerations of distributive justice are frequently ignored altogether, or viewed as subsidiary, in these works, basically for reasons of analytic method (pertaining principally to their definition of the object of study). Many of them follow the utilitarian tradition of this field, originating with Beccaria (1764), often combined with a strict version of the behavioral assumption of *homo economicus* (see Posner (1981), and Kaplow and Shavell (2002), for comprehensive presentations and discussions of numerous examples, including some that do not fit the simplistic generalization offered here). The position of *method* that typifies much of this body of studies must not be confused with the *normative* position of such global theories of justice as those of Hayek (1974) or Nozick (1974). These two writers' rejection of contemporary redistributive welfare states is a consequence of their particular appreciation of the public and private modes of action within their norm of justice. We return to Hayek's and Nozick's theories below (in section 4.1, footnote 22).

<sup>15</sup> See the chapters by Fong, Bowles and Gintis (Chap. 23), Blanchet and Fleurbaey (Chap. 24), and Cigno (Chap. 25) of Kolm and Mercier Ythier (2006) for comprehensive overviews of the positive and normative (including motivational) foundations of the institutions of the welfare state.



decision to become, in themselves, issues of justice.<sup>16</sup> Dworkin's theory is a nice case of the deduction of a norm of distributive justice from the norm of commutative justice: the norm of communication defined by the latter (that is, perfect competition<sup>17</sup>) substitutes for the unanimity of individual views, lost following the partial removal of the veil of ignorance. Kolm's *Liberal Social Contract* (1985) and Sen's *Idea of Justice* (2009), both considered below, develop theories of benevolent justice from two different norms of rational, communicative decision making within society. Both of their norms are inherited from the Enlightenment: Kolm draws on Locke's variant of the social contract (which may be viewed itself as an Enlightenment's extension of Aristotle's commutative justice) while Sen derives his norms from the impartial spectator of Adam Smith. We briefly sketch these approaches below (in sections 4.1 and 4.2), and then complement this outline with a brief consideration of related constructs relative to benevolent reciprocity (4.3) and the ethics of caring (4.4).

#### 4.1-Benevolent justice from extended commutative justice<sup>18</sup>

Kolm's liberal social contract (1985) employs the same constructive device as Dworkin, pushed to the limit. It removes the veil of ignorance completely, and substitutes for it the ideal norm of communication of the liberal social contract, which includes perfect competition as a subset of this ideal norm of communication in and through market exchange. The (distributive) liberal social contract is defined from a hypothetical initial position of *perfect contracting*, characterized *lato sensu* as follows: (i) individuals know their own characteristics, that is, their preferences relative to their own consumption and relative to the distribution of private wealth and welfare in society, and their "initial" endowments of market commodities (i.e. their individual endowment prior to redistribution under the social contract); (ii) they can engage freely in the exchange of commodities in a complete system of competitive markets ; (iii) they can engage freely in redistributive wealth transfers that are unanimously agreed upon (with indifference by other individuals being understood as agreement); (iv) the redistributive social contracts referred to in point (iii) are established in the same conditions as the market contracts referred to in point (ii), that is, the related costs of information, transaction and enforcement are null; (v) the redistributive transfers of the social contract are lump-sum, that is, they involve no distortion in competitive market prices whatsoever. The initial position *stricto sensu* consists of a definite putative (that is, virtual) competitive equilibrium associated with the initial distribution of endowments. It is, in other words, an imaginary position of competitive market equilibrium *prior* to redistribution under the social contract.<sup>19</sup>

The distributive liberal social contract consists of a distribution of individual endowments and of an allocation of resources such that: (i) the allocation of the social contract is a competitive market equilibrium relative to the endowment distribution of the social contract; (ii) the

---

<sup>16</sup> And it justifies, more generally, the development of an ethics of communication. This was achieved by K.-O. Apel (1987) and J. Habermas (1990). See also Ricoeur (1992), 9, pp. 280-283.

<sup>17</sup> See Mercier Ythier (2013) for developments on the familiar interpretation of competitive exchange as commutative justice and their common origins in the *Nichomachean Ethics*.

<sup>18</sup> The account of the liberal social contract given here is based on the version developed in Mercier Ythier (2010 and 2011).

<sup>19</sup> The logical time sequence implied by this definition of the hypothetical initial position admits a natural interpretation in terms of the historical time sequence of the economic revolution of the late eighteenth century and the welfare state revolution of the late nineteenth century: see Mercier Ythier (2011), 1, notably footnotes 4, 5 and 7, for an analysis of the various aspects of the notion of time involved here. See also footnote 20 below on another aspect of the same subject.

allocation of the social contract is unanimously (weakly) preferred, in terms of individuals' distributive preferences, to the competitive market equilibrium of the initial position; and (iii) the allocation of the social contract is a (strong) Pareto optimum relative to individuals' distributive preferences. To restate these three points, it is *a distributive (Pareto) optimum unanimously (weakly) preferred, in terms of distributive preferences, to the (competitive) market equilibrium of the initial position.*

In previous publications I have established the logical robustness of this notion by showing the existence of a solution (Mercier Ythier, 2011, 4.1, Theorem 1), and the existence of a non-trivial set of solutions for a sufficiently rich set of general social systems (Mercier Ythier, 2010, 5.1, Theorem 2).<sup>20</sup> These works also establish relevant connections between social contract and the social welfare function of Bergson-Samuelson (Mercier Ythier, 2011, 4.2, Theorem 2), and also between social contract and Lindahl pricing (Mercier Ythier, 2010, 4, Theorem 1).<sup>21</sup>

The norm of distributive justice so defined is compatible a priori with any degree of wealth redistribution, provided that it be unanimously wished, or at least wished by some individuals and vetoed by none. The modalities of its implementation depend on the particular mix of contract and public failures confronting practical redistribution (and assumed away for the derivation of the norm). The distinction between “contractual” and “public” activities is pointless as long as one considers the norm of the social contract, that is, as long as one assumes perfect contracting, because government activities (if any) are then but one of the possible expressions of the social contract. The same distinction does however become relevant as soon as one considers the implementation of the norm, because informational and behavioral contingencies (“imperfections”) condition implementation, and because these contingencies involve systematic differences between governmental and nongovernmental (e.g. individual or associative) modes of action. Lump-sum redistributive taxation may, for example, prove to be an adequate substitute for charitable redistribution if the latter involves significant free-riding behavior. Conversely, charitable redistribution may serve as an adequate substitute for redistributive taxation if taxation involves significant distortions, such as significant labor disincentives for example. That is, the “public” is simply a part of the “contract” in the ideal conditions of the norm of justice, while the public failures (or imperfections) and the contract failures (or imperfections) are distinct conditions of the implementation of the norm of justice in the positive (i.e. “imperfect” or “second best”) conditions of practice.<sup>22</sup>

---

<sup>20</sup> The latter theorem shows that, in the situations where there exists a unanimous agreement for some effective redistribution, the social contract solutions form, typically, a set of dimension  $n-1$ , where  $n$  is the number of individuals. This indeterminacy of the set of social contract solutions abstractly defined above is reconciled with the determinacy (that is, the uniqueness) of the liberal social contracts of Kolm or Nozick by adopting a historical or *ex post* view of the social contract, as suggested notably by Nozick (1974, II, 7, pp. 153-155). There are, typically, infinitely many social contract solutions *ex ante*, that is, *before* social contract redistribution, but there remains only a single one *ex post*, that is, *after* social contract redistribution, because an allocation has been chosen then by definition, and because the iterative application of the same norms of justice to the social contract allocation now yields the status quo. More precisely, the social contract allocation is a distributive optimum by definition, so that the set of market equilibria that are unanimously weakly preferred to it reduces to the social contract allocation itself: that is, the solution set becomes a singleton (the status quo allocation) if a new social contract is established after social contract redistribution has been performed.

<sup>21</sup> The latter results elicit some consequentialist aspects of the norm of the liberal social contract, which relate it closely to the welfare economics and the public economics of Paul Samuelson. See also the following footnote.

<sup>22</sup> The main cases of contract failures discussed in the economic literature are market failures, and the coordination (“free-riding”) issues associated with voluntary redistributive transfers. The main cases of public failures usually considered in public economics stem, on the other hand, from the costs and issues of the

Note, finally, that the attractiveness of commutative justice as a norm of justice goes well beyond modern liberalism. There is, of course, the reference by Aristotle, and we detailed above the role played by perfect competitive exchange in the construction of Dworkin's scheme of fundamental insurance.<sup>23</sup> We may add Emile Durkheim to the list. The norm of justice outlined in one of the last chapters of his *Division du Travail Social* (1930, III, 2) is simply a form of commutative justice, extended to all social activities.

#### 4.2-Benevolent justice from the benevolent impartial spectator of Adam Smith

Sen's *Idea of Justice* (2009) develops the non-consequentialist potentialities of the sympathetic impartial spectator of Adam Smith.<sup>24</sup>

The informational and motivational bases of the construct are jointly provided by Sen's concept of capabilities. This notion refers to the multiple dimensions along which human persons can develop (i.e. achieve improvements for and/or of themselves) such as education and training activities, health, housing, labor activities, individual welfare, or individual rights and liberties. We can briefly compare Sen's construct with the informational and motivational bases of the four theories of justice already discussed above. Sen's capabilities differ from Rawls's primary goods on three important grounds at least. First, they include the personal characteristics (circumstances and preferences) of individuals. Second, on epistemic and interpretive grounds, Rawls's Kantian references to the nature of human beings as free and equal rational beings are viewed as irrelevant because, as metaphysical conceptions, they are not accessible to scientific knowledge (that is, their truth or falsity cannot be established by existing scientific methods). Third, Sen qualifies the lexicographic priority attributed to rights and liberties by Rawls's reflective equilibrium (it is denied as a categorical imperative, although with nuance and precaution: see Sen, 2009, III, 14, pp 299-301). Sen's capabilities include Dworkin's extended resources and individual personal preferences in their account of

---

administration of public action, from distortionary taxation, and from the problems of identification ("revelation") of individuals' preferences relative to public goods. The positions adopted by Nozick (1974) and Hayek (1974) appear quite extreme in these respects. The former denies any practical importance to contract failures, even in matters of voluntary redistribution (1974, II, 8; see especially the subsection relative to philanthropy, pp. 265-268), while the latter views the contemporary expansion of welfare states as threatening dramatically the fragile spontaneous order of developed market economies and societies (a case of fundamental public failure, therefore, since public action is supposed to create the conditions of the collapse of the market economy and of the institutions of liberal political democracy). Their similar conclusion that public redistributive transfers should be kept to a minimum (Nozick, 1974, III, pp. 333-335; Hayek, 1974, I, 6, pp. 141-143) follows, therefore, from their appreciation of the conditions of practice, that is, from positive statements (which may be true or false) relative to the implementation of the norm of justice of the liberal social contract. Long-run historical experience, in Western Europe and elsewhere, does not, in my opinion, support such (categorical) practical views.

<sup>23</sup> Dworkin's fundamental insurance is not a liberal social contract notably because it derives its norm of justice from hypothetical preferences of individuals, that is, the individual expected utilities constructed under his hypothetical veil of ignorance (see footnote 12 above). The liberal social contract deduces the norm of justice from actual individual preferences. The same remark applies to Rawls's social contract. Veil-of-ignorance theories of justice, such as those of Dworkin, Rawls or Harsanyi, impose norms of justice on individual preferences and on various other individual characteristics (the list of which depends on the precise content of what is hidden under the veil). The liberal social contract is *positive* relative to individual preferences and individual endowments of market commodities; it is *normative* relative to the conditions of social communication and exchange; it is *jointly positive and normative* relative to the practical conditions of implementation of the norm, in the sense that it identifies and assesses the conditions of practice by appreciating their distance from the norm (that is, the "imperfections" of the real conditions of practice are defined and evaluated relative to the ideal conditions of perfect contracting).

<sup>24</sup> See Mercier Ythier (2013, 2.1), for further discussion of the ethics of Adam Smith.

the ethically relevant characteristics of individuals. They also encompass the informational and motivational bases of the distributive liberal social contract (individuals' ordinal preferences relative to own consumption and the distribution of wealth, and their endowments in market commodities) and the informational basis of utilitarianism (individuals' objective or cardinal utilities, if any), but are much wider in scope than these. In particular, Sen's *Idea of Justice* is not utilitarian, obviously, nor is it consequentialist either, since individual liberties are major items on the list of capabilities. Among the informational bases of the five theories of justice considered in this section, Sen's is clearly the widest.

There is no notion of a hypothetical veil of ignorance, whether thin or thick, in Sen's construct, nor is there any reference, more generally, to hypothetical conditions constituting an "original" or "initial" or "ideal" position or standpoint for the definition of the norm of justice. Such hypothetical conditions are constitutive of what Sen calls the *transcendental* approaches to justice, and which he opposes to his own, *comparative* approach (2009, pp. 5-27).<sup>25</sup> The construction of comparative solutions of justice relative to the capabilities relies on social choice theory and voice, construed as two complementary aspects of the same process (2009, I, 4). *Social choice theory* refers to the set of methods elaborated in order to produce, from the preferences of the individuals concerned, a ranking of alternatives in a set of accessible social realizations, which respects a number of selected requirements of rationality and ethics. It establishes a methodological framework for the second aspect of the process, *voice*, which refers to public reasoning and public deliberation relative to solutions of justice, and particularly to the elaboration of public reasoning through public deliberation about matters of justice. The solutions of justice, so elaborated, depend on the context of choice; and they need not be determinate in a given context, that is, the social ranking may remain incomplete, either for contextual or for more essential reasons (Sen, 2009, 4, pp. 106-111). The reference to Smith's impartial spectator (e.g. Sen, 2009, I, 1, pp. 44-46) points to the main procedural feature of the method, namely, the requirement that the voice of any person concerned be heard and taken into account, in relevant conditions of benevolent impartiality, in the process of elaboration of the solutions of justice. This feature establishes an essential difference between Sen's construct on one hand, and the social contract theories of justice on the other hand, since the latter take into account only the preferences of the individuals who belong to the political unit constituted by the contract or covered by it. In other words, Sen's notion of justice (like utilitarianism) neither supposes nor implies the political institution of a state.

#### **4.3-Benevolent reciprocity, the good society and the possibility of justice**

Reciprocity has been a basic topic of economic anthropology since (at least) Mauss's classical *Essai sur le Don* (1924).<sup>26</sup> Construed as gift-exchange, and carefully distinguished from

---

<sup>25</sup> Sen's rejection of definite ("transcendental") norms of justice, and the extensive scope and comprehensive contents of his notion of capabilities, are related features of the construct. Their combination may result, sometimes, in a weakening of the distinction between ethical norms and social facts, or between ethical assertions and descriptions of social reality. Smith's ethical theory has analogous characteristics: its eclecticism relative to the bases of moral approval, combined with the psychological nature of its process of empathetic construction of the moral self, makes Smith's work sometimes very close to pure psychological theory or to the fine psychological analyses of literary moralists. In fact, I consider Smith's comparison of pride and vanity (1759, VI, Section III) to be a masterwork of literary moralism. These characteristics go together with Smith's propensity to formulate the demands of normative ethics directly, as deontological requirements (calls to the sense of duty), which may then seem sometimes artificial or stiff, as lacking sufficient reasons *inside* his ethical theory itself: norms must be *superimposed*, if they are not *derived*.

<sup>26</sup> See, for example, the historical overview by Hann (2006, 2).

contract, it is interpreted as a form of communicative action that expresses and by the same token actualizes the social bond that cements human communities (e.g. Godelier, 1996, 2000; Sacco et al., 2006, 2).

Economic anthropology grasps reciprocity as a set of (conceptually related) facts of social life. These facts are analyzed as complex arrangements of individual motives and behavioral norms, the presentation of which far exceeds the scope of the present study.<sup>27</sup> We want to concentrate here, instead, on the ethical norm of reciprocity, and on the articulation of this norm in relation to both benevolence and justice.

Mauss himself ended his anthropological analysis of the practices of gift-exchange of the Maori *kula* and of the Kwakiutl *potlatch* with moralistic conclusions in which he deplored the lack (actually, in his evolutionary perspective, the loss), in contemporary developed market economies, of the climate of generosity and affective warmth of the large scale gift-exchanges of these other (“archaic”) societies. Polanyi’s *Great Transformation* (1944) systematized, on both the analytical and the ethical side, this critical perspective on the substitution of market for reciprocity in the course of long-term development. Kolm’s *Bonne Economie* (1984), and the notion of civil happiness presented by Bruni and Zamagni (2007), are more recent examples of related conceptions. These authors view benevolent reciprocity as a norm for the ethical evaluation of the functioning of contemporary economies, and also as a practical remedy for some of its major deficiencies. The main such deficiency cited is an ethical failure, which follows from a constitutive lack of benevolent feelings in the human relationships conveyed by the anonymous (“non-tuistic”) interactions of individuals through efficient markets, and also by the impersonal modes of action of the public administrations of impartial states of law. An economy of benevolent reciprocity already exists, it is argued, inside the wide set of activities that are not covered by market or state institutions. Its expansion, absolute and relative, is advocated as the solution to the ethical failure above, and also as a practical solution to other, more standard cases of market or public failure.<sup>28</sup>

The contributions mentioned above are works from the social sciences. An accomplished philosophical discussion of benevolent reciprocity as ethical norm may be found in the ethics of Paul Ricoeur (1990a, 1990b, 1990c), and particularly in the studies 7 through 9 of his *Oneself as Another* (Ricoeur, 1992).<sup>29</sup> Ricoeur summarizes his ethics in a single formula, of

---

<sup>27</sup> Important recent developments on these subjects may be found in Elster (2006; 2009, 5; 2011).

<sup>28</sup> One notes interesting differences in the detail of the discussion of standard (“objective”) public and market failures in these two works, related to significant differences in the historical context in which they were written. Kolm’s book was written a few years before the final collapse of the planned economies and authoritarian political systems of Eastern Europe, which followed the fall of the Berlin wall in 1989. The book devotes, accordingly, significant space to the problems of centrally planned economies. Bruni and Zamagni, meanwhile, devote great attention to the problems of the welfare states and labor markets of Western Europe—problems that were only emerging in the early 1980s and gradually gained importance over the subsequent 20 years, finally crystallizing in the current debt crisis. The two books thus stand, historically, at the beginning and the end of a consistent period of three decades of contemporary economic history, one that stands in sharp contrast to the first three decades of the post-war period with regard to the efficacy of public economic action. The important structural changes that characterize this historical period provide the factual justifications for the type of question studied in these works, namely the comparative evaluation of global economic systems (such as market, reciprocity, planning, and their various combinations) on the basis of their ethical and other practical characteristics. This comparison is conceptualized by Kolm through individual preferences relative to the modes of transfer of scarce resources (1984, 15, notably pp. 418-419), and by Bruni and Zamagni through the notion of relational goods (2007, 9, notably pp. 239-244).

<sup>29</sup> These three studies develop Ricoeur’s ethics, which is constructed from a theory of action based on the person as a singular, communicative being. A concise, tractable version of this (sophisticated) ethical theory is presented in his short essay *Ethique et Morale* (1990b). The study *Amour et Justice* (1990c) considers the

Aristotelian tone: “*aiming at the “good life”, with and for others, in just institutions*” (Ricoeur, 1992, p. 172; emphasis in original). Benevolent reciprocity corresponds to the second of the three “moments” of this overall ethical perspective, namely, the moment of interpersonal relations (“with and for others”). It will suffice, for the limited purposes of the present study, to enumerate the main structural features of this aspect of Ricoeur’s construct. They consist of *solicitude*, *reciprocity*, and the *Golden Rule*. Let us examine them in turn. Solicitude is an individual motive, corresponding to an affect (feeling) of spontaneous, benevolent sympathy. Solicitude stems from a fundamental partition (“unfolding”) of Self into Same and Other, which notably involves: (i) the (felt) recognition by (my)self of the (radical) otherness of other than (my)self; (ii) the sentiment of need and lack of the otherness of the “other self”, such as what we experience in friendship; and (iii) the balancing of this sentiment of need and lack by a feeling of benevolent sympathy, which expresses the (felt) recognition that the other self is *oneself as myself*. Reciprocity is shared solicitude; it achieves the exact equivalence of “the esteem of the *other as a oneself* and the esteem of *oneself as an other*” (Ricoeur, 1992, p. 194; emphasis in original). The Golden Rule, finally, is the deontological norm that makes benevolent reciprocity a duty. Ricoeur provides three classical formulations of this norm of reciprocity, from the Jewish and Christian traditions (1992, p. 219): “Do not do unto your neighbor what you would hate him to do to you. This is the entire law; the rest is commentary” (Babylonian Talmud, *Sabbath*, 31a); “Treat others as you would like them to treat you” (Luke 6:31); and “Love your neighbor as yourself” (Leviticus 19:18 ; Matthew 22:39).

Following the set of analyses above, benevolent reciprocity appears best construed as the necessary intermediary between the psychological fact of altruism and the norm and institutions of justice. Altruism and benevolent reciprocity are necessary, first, for the achievement of the best potentialities of human beings as singular, communicative persons (i.e., for the achievement of the good life in the good society). Related psychological facts, and norms of benevolent reciprocity as pervasive social facts, are accessible to systematic evaluation by the social sciences and by the rapidly expanding neurosciences. Elster’s works (2006, 2009, 2011) are significant recent contributions to this scientific program. Solicitude and the Golden Rule substitute, in Ricoeur’s ethics, for Kant’s famous (metaphysical) “fact of reason”, namely, the nature of human persons as “free and equal rational beings with a liberty to choose”, to use Rawls’s words. The universalization of the Golden Rule under the heading of Kant’s second categorical imperative, that is, the respect of persons as universal obligation, is, second, the necessary step that submits justice to the ethical aim in Ricoeur’s construct. It corresponds, practically, to the ban on violence and on physical abuse in human interactions, and notably in the political process of elaboration of just institutions. Altruism therefore appears, in this latter respect, necessary also for submitting justice to the requirements of practical wisdom, and hence for an overall consistent articulation of justice, prudence (i.e., practical wisdom), and benevolence in a unified, humanist ethical perspective that has practical usefulness.<sup>30</sup>

---

economy of gift-giving (love, reciprocity and the Golden Rule) and its relation to justice from a theological perspective. I am not aware of the existence of English translations of the latter two texts.

<sup>30</sup> An alternative unified humanist ethical perspective on justice may be found in Dworkin’s beautiful *Justice for Hedgehogs* (2011). This work is Kantian throughout, constructed on the key values of dignity, responsibility and truth, and consistent with the individualistic features of his earlier contributions. Nevertheless, I find that the anthropological foundations of Ricoeur’s ethics (drawing from Levinas, Arendt and many others including Kant) are more accurate, comprehensive and robust than those of Kant and, consequently, of Dworkin or Rawls. Stated in another way, Kantian anthropology is too individualistic (that is, it relies too exclusively on individual autonomy) to be fully convincing in my view, when compared with Ricoeur’s. I return to this point in section 5.2.1.

#### 4.4-Justice from (partial) benevolence: Slote's ethics of caring

Slote's ethic of virtue (2001) shares important common features with Ricoeur's ethics, but has a simpler, monistic structure. For Slote, the sole basis of ethical appraisal is the motive of benevolence, construed as a spontaneous feeling of sympathy that urges the benevolent person to care for the welfare of the beloved.

Slote considers three conceivable variants of benevolent motive as possible candidates for his definition of moral excellence: *caring*, as defined in the original contributions of Carol Gilligan (1982) and Nel Noddings (1984); *universal benevolence*; and *universal love*. Caring, or benevolent feeling in *interpersonal* relations, is synonymous with Ricoeur's solicitude. Universal (humanitarian) benevolence is construed, by contrast, as impersonal ("aggregative") benevolence, that is, it views the beneficiaries of benevolence as a fungible set of anonymous, substitutable individuals. Universal love, finally, is (ideal) Christian love, corresponding to the universalization of Ricoeur's third variant of the Golden Rule ("Love your neighbor as yourself"). Slote (2001, 5) selects caring as his definition of moral excellence.

Caring is a partial form of benevolence, in the sense that the caring person is emotionally involved with the individuals with whom he or she cares, and with them only. This constitutive partiality of the caring motive raises obvious questions about the possibility of deriving a valid conception of justice within the ethics of caring, or at least within such a consistent, monistic, agent-based ethics of caring as Slote's. This issue is addressed in two steps.

Slote first introduces a notion of *balanced* caring (2001, 3, section 2). He notices that caring itself, as concern for the near and dear, can coexist, in a single person, with humanitarian benevolence (also called, by extension, humanitarian caring), or concern for the welfare of any unknown person. A morally good or decent person, Slote argues, spontaneously establishes balanced caring "between the concern she has for her intimates *considered as a class* and the concern she has for all (other) human beings *considered as a class*" (2001, p. 70; emphasis in original). This balance should not be construed, he adds, as a quantitative balance between the intimate and the humanitarian caring motives, but, rather, in the manner of Aristotle's generic characterization of virtues, as a qualitative middle term between two extremes, namely, the extreme of exclusive intimate caring on the one hand and the extreme of universal (impartial) benevolence on the other hand. Balanced caring establishes, in other words, some sort of qualitative equivalence (and "communication") between the intimate caring and the humanitarian caring motives of the caring person.

The second step of Slote's derivation of justice from his agent-based ethics of caring consists of a move from the interpersonal to the institutional (political, social, judiciary) plane (2001, 4). The equivalent, on the institutional plane, of concern for the near and dear is the caring person's benevolent concern for the good of his or her country. Social (national) justice is then construed as a state of public affairs where "national public institutions and laws reflect (enough) such concern on the part of (enough of) those who create (or implement or maintain) them or at least don't reflect a (great) lack or deficiency of this motive" (2001, p. 100). Likewise, the equivalent, on the institutional plane, of the humane concern for unknown others is the benevolent concern of the caring person for the welfare of other nations. Slote argues that balanced caring, when transposed to relevant public institutions and laws in the

same way as caring was transposed to concern for one's country, provides the appropriate motive for a valid notion of justice in international public affairs.

## **5-Some practical conclusions: Distributive justice, benevolence and the welfare state**

In this section, we return first to some basic features of this analytic review of theories of distributive justice. Next we offer an exposition of the reasons why, in our view, justice in distribution must be supported by the virtues of benevolence and prudence in order to produce sustainable, practical distributive justice; we also suggest how this support should be provided. We conclude by characterizing contemporary welfare states as practical realizations of the norm of distributive justice of the liberal social contract.

### **5.1-Distributive justice as a field**

The main theories of distributive justice considered above occupy complementary positions in their field, which may be characterized, respectively, as speculative-deontological (Kant-Rawls), speculative-consequential (Hume-Harsanyi), practical-constitutional (Locke-Kolm), and pragmatic (Smith-Sen).<sup>31</sup>

Selecting a particular theory in this menu consists not so much, in our view, of selecting one's preferred substantive conception of distributive justice as of selecting a particular perspective on the field as a whole, which, per se, need not imply any substantive conception whatsoever. The speculative perspective emphasizes conceptual clarification, logical consistency and generality, while the practical perspective puts the main emphasis on practical applicability. Rule-based (i.e., deontological or constitutional) perspectives emphasize ethical norms, while consequential and pragmatic perspectives put the main emphasis on practical achievements.

Considered from a substantive point of view, the four positions outlined above participate in liberal ethical theory, characterized notably by (i) the analytic identification of liberty with human development, and (ii) the favorable normative evaluation of human development in concept (duly balanced through the careful consideration of the pros and cons of actual human development).

The object of the present section is to provide some *overall* guidance for *practical* distributive justice. We adopt, accordingly, the practical-constitutional viewpoint in this section, that is, the perspective of the liberal social contract.

### **5.2-Benevolence, the social contract, and the welfare state**

The welfare state refers to the set of public rules and actions designed to achieve a suitable distribution of wealth and welfare in a political society. It *is* practical distributive justice. An essential component of the welfare state is the set of social insurances that provide for such needs as retirement income and health care. Veil-of-ignorance theories of distributive justice endorse this fact, implicitly at least, by grounding their norms of redistribution on extended

---

<sup>31</sup> The reader will notice the absence of Hobbes in this enumeration. The object of Hobbes's social contract theory is the very existence of the polity, that is, the conditions for the prevalence of peace over war in a political society. These conditions are necessary for considerations of distributive justice to make any sense, in either a domestic or an international context. We take them for granted implicitly in this paper, that is, we restrict ourselves to examining political societies that are already, substantially, in a state of peace. The main extension of the Hobbesian tradition in contemporary economics is James Buchanan's public choice theory (1962, 1975).



insurance motives of individuals, that is, on rational individual response to radical uncertainty relative to personal identity (Harsanyi), access to primary goods (Rawls), or personal talents and handicaps (Dworkin). We briefly examine the institutions of the welfare state below, regarding their foundations (5.2.1) and methods (5.2.2), as parts of a *practical* distributive liberal social contract.

### **5.2.1-Benevolence and the welfare state**

Rawls notes accurately that the utilitarian norm fails to “take seriously the distinction between persons” (1971, p. 27). I will argue here, in turn: (i) that the Kantian norms of distributive justice of Rawls and Dworkin fall short of a satisfactory ethical treatment of the *dependent* person; (ii) that benevolent caring is the appropriate norm for the latter; and (iii) that this norm should take priority over social insurance motives, however defined. I conclude with a brief recapitulation of the main motives involved in the social culture that supports the institutions of the welfare state.

The standard of evaluation of Kantian ethical norms is individual autonomy. The difficulties that may result from use of that standard for an appropriate ethical treatment of dependence may be clarified through the following two arguments, the first one dealing with situations of extreme dependence and the second one with more ordinary circumstances. Consider, first, the case of an individual who so much values autonomy that he or she decides, after careful consideration of the pros and cons, that he or she would rather commit suicide than fall into some state of wretched dependence. That is, life is worth living only if this person enjoys sufficient personal autonomy. The autonomous self denies existence, actively, to the other, prospective, dependent self. The two cannot coexist in the same ethical world. Now let us consider less extreme situations such as aging persons experiencing the early stages of Alzheimer disease. The Kantian response to such standard situations of dependence will normally consist of large transfers of resources in order to, first, maintain or restore the personal autonomy of the beneficiary as much as possible, and, second, compensate for irreversible losses of autonomy. In many practical circumstances, however, the full restoration of autonomy is impossible, and the compensating transfer implies the coexistence, in this ethical world, of two unequal perspectives—namely, the perspective of the norm and the perspective of the dependent, the latter depreciated by the former.

The issues of the impossible or depreciated existence of the dependent have significant practical content in contemporary welfare states. The dependence of the aged, notably, is becoming a major practical issue of distributive justice, due to the demographic consequences of advanced development. The risk of dependence is gradually emerging as one of the main dimensions of social insurance, and laws governing assisted suicide and euthanasia are increasingly becoming matters of public debate. Benevolent caring provides the ethical framework for an appropriate treatment of these questions in their context. The ethics of caring was developed initially by its promoters in the early 1980s as the ethics that is most suitable for use in educating young children. Its main characteristic, which is also the source of its effectiveness, consists of including the perspective of the dependent inside the perspective of the benevolent caring person, while supporting the former’s personal identity (i.e., the individual singularity of the dependent). It is the solution of principle—by definition, so to speak—of the ethical issue raised above.

Benevolent caring is a form of personal altruism. The latter may be construed, more generally, as the fact of individual moral conscience that consists simultaneously of (i) the felt

recognition of the existence of others, and (ii) the positive sentiments associated with it (that is, one senses the existence of people other than oneself as something good). The norms of altruism stem from the recognition that altruism may fail, with potentially tragic consequences, when one actively denies the existence of others.<sup>32</sup> The norms of altruism ban the active denials of human sociality, and should have priority over all other ethical norms—that is, in Aristotelian terms, over all norms pertaining to the good life of human beings as social animals. This general line of ethical argument, applied to distributive justice, implies the priority of the norm of benevolent caring over the other ethical norms of that field, whether Kantian, utilitarian or otherwise. That is, the conception of distributive justice that we defend here implies that the norm of benevolent caring should prevail whenever it conflicts with another norm of distributive justice.

A welfare state is a practical solution of the coexistence of individual perspectives in a political society, considered from the standpoint of individual welfare. It is supported by its various participants—e.g. the taxpaying citizens, the beneficiaries of welfare services, and the government, administration and staff—for various motives, which include notably the motives of caring and reciprocity, deontological motives relative to the individual rights concerned, and consequential motives relative to individuals' welfare. The set of motives particular to the participants in a welfare state make up the social culture of their political society. The institutions of a welfare state prosper, decay, or are revived, in parts or as a whole, in relation to this culture.<sup>33</sup>

### 5.2.2-The social contract and the welfare state

The method for achieving distributive justice in the practical-constitutional perspective considered here is the social contract. The contract consists of two main parts: market equilibrium, and unanimous agreement regarding public action. Agreement relative to public action, in particular, is reached through public deliberation (“voice”: see section 4.2 above), which means in practice both (i) that support to public action is loud and determinate enough to be durably heard in public debate (voiced agreement), and (ii) that opposition to public action is either not loud enough or not determinate enough to be durably heard in public debate (lack of voiced disagreement).

The practical limits of agreement are conventionally designated as its “failures” (see section 4.1 above). They are places of confrontation between the contract and adverse social facts, which provide the basis for the prudential assessment of the practical sustainability of the norm of justice of the social contract. *Market failures* consist principally of externalities and the public good problem, and the incompleteness of capital markets. The latter, notably, results in radical uncertainty, as illustrated by the role of the US financial system at the starting point of the global economic crises of 1929 and 2007. *Government failures* follow from the large costs of determining, implementing, and enforcing public action. They result in the limited ability of public administration to address the variety and variability of the circumstances that determine its effectiveness. Recent characteristic examples were provided by the collapse of the political societies of Eastern Europe after 1989 and, to a lesser extent,

---

<sup>32</sup> The Golden Rule and Kant's second categorical imperative are norms of universal altruism.

<sup>33</sup> One may wish to construe practical distributive justice as one of Walzer's (1983) spheres of justice, or a subset of them. I prefer, nevertheless, to characterize it as a *perspective* of justice, consistently with the analytic, relativist epistemology that runs throughout the present article. There is, within this latter view, no such relevant *object* as justice or distributive justice, but, rather, a variety of perspectives of justice, which interfere with each other, combine, and succeed to the extent of their evolution in actual institutions and cultures.

by the persistent difficulties of Western European welfare states in the recent decades and by the euro crisis of 2011. Another limitation common to both market and public agreements was referred to above as their *ethical failure* (section 4.3). This type of failure follows from the impersonality of market and government as modes of mass communication and action. The interpersonal relations required to sustain the existence and development of human beings as singular, communicative persons are supported by a variety of institutions including the family and other nonprofit, nongovernmental organizations. Such institutions constitute the economy of gift-giving. This gift economy is maintained in its relative position and share by its restricted access to the main sources of capital accumulation, namely, profit and tax.

This set of practical limits shapes a Coase equilibrium made of three irreducible components: market, voice, and the gift. The relative shares of these three components may vary substantially according to place and time.<sup>34</sup> The practical distributive liberal social contract consists of any sustainable realization of such a Coase equilibrium, in the definite conditions of a political society.

## References

- Apel, K.-O. (1987). *Sur le problème d'une fondation rationnelle de l'éthique à l'âge de la science : L'a priori de la communauté communicationnelle et les fondements de l'éthique*. Lille : Presses Universitaires de Lille.
- Aristotle (1941). *Ethica Nichomachea*. English translation by W.D. Ross, in *The Basic Works of Aristotle*, ed. Richard McKeon. New York: Random House.
- Bruni, L. and S. Zamagni (2007). *Civil Economy: Efficiency, Equity, Public Happiness*. Bern: Peter Lang AG, International Academic Publishers.
- Buchanan, J. M. and G. Tullock (1962). *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. IN: Liberty Fund, Inc.
- Buchanan, J. M. (1975). *The Limits of Liberty: Between Anarchy and Leviathan*. IN: Liberty Fund, Inc.
- Durkheim, E. (1930). *De la Division du Travail Social*. Paris : Presses Universitaires de France.
- Dworkin, R. (1981a). What is equality? Part 1: Equality of welfare. *Philosophy and Public Affairs* 10: 185-246.
- Dworkin, R. (1981b). What is equality? Part 2: Equality of resources. *Philosophy and Public Affairs* 10: 283-345.
- Dworkin, R. (2011). *Justice for Hedgehogs*. Cambridge (MA) and London: The Belknap Press of Harvard University Press.
- Elster, J. (2006). Altruistic behaviour and altruistic motivations. Chap. 3 in S.-C. Kolm and J. Mercier Ythier eds., *Handbook of the Economics of Giving, Altruism and Reciprocity*, (Amsterdam : North-Holland) 183-206.
- Elster, J. (2009). *Le Désintéressement. Traité Critique de l'Homme Economique I*. Paris : Seuil.
- Elster, J. (2011). Reciprocity and norms. In M. Fleurbaey, M. Salles and J.A. Weymark eds., *Social Ethics and Normative Economics. Essays in honor of Serge-Christophe Kolm* (Berlin, Heidelberg: Springer-Verlag) 303-26.
- Fehr, E. and K. M. Schmidt (2006). The economics of fairness, reciprocity and altruism-Experimental evidence and new theories. Chap. 8 in S.-C. Kolm and J. Mercier Ythier eds., *Handbook of the Economics of Giving, Altruism and Reciprocity*, (Amsterdam : North-Holland) 615-94.
- Fleurbaey, M. (1996). *Théories Economiques de la Justice*. Paris : Economica.
- Gilligan, C. (1982). *In a Different Voice: Psychological Theory and Women's Development*. Cambridge, MA : Harvard University Press.
- Godelier, M. (1996). *L'Enigme du Don*. Paris: Fayard. Translated as *The Enigma of the Gift*, Cambridge: Polity Press (1999).
- Godelier, M. (2000). Things you don't give or sell but which you keep: Valuable and social objects. Chap. 9 in L.-A. Gérard-Varet, S.-C. Kolm and J. Mercier Ythier eds., *The Economics of Reciprocity, Giving and Altruism*, London: Macmillan.
- Grant, S., A. Kajii, B. Polak and Z. Safra (2010). Generalized utilitarianism and Harsanyi's impartial observer theorem. *Econometrica* 78: 1939-71.
- Habermas, J. (1990). *Moral Consciousness and Communicative Action*. Cambridge: MIT Press.
- Hann, C. (2006). The gift and reciprocity: Perspectives from economic anthropology. Chap. 4 in S.-C. Kolm and J. Mercier Ythier eds., *Handbook of the Economics of Giving, Altruism and Reciprocity*, (Amsterdam: North-Holland) 208-23.

---

<sup>34</sup> The nonprofit sector (i.e., charities, cooperative firms, and mutual insurance companies) accounts for 10 percent of French GDP currently, for instance. This share could climb to 13% or thereabouts in the moderately near future as a partial response to current issues in public finance.

- Harsanyi, J.C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy* 61:434-35.
- Harsanyi, J.C. (1955). Cardinal welfare, individualistic ethic, and interpersonal comparison of utility. *Journal of Political Economy* 63:309-21.
- Harsanyi, J.C. (1977). *Rational behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.
- Hayek, F. A. (1974). *Law, Legislation and Liberty*, vol. 1. Chicago: University of Chicago Press.
- Kaplow, L. and S. Shavell (2002). *Fairness versus Welfare*. Cambridge, M.A.: Harvard University Press.
- Kolm, S.-Ch. (1985). *Le Contrat Social Libéral*. Paris : Presses Universitaires de France.
- Kolm, S.-Ch and J. Mercier Ythier (2006). *Handbook of the Economics of Giving, Altruism and Reciprocity*. Amsterdam : North-Holland.
- MacIntyre, A. (1981). *After Virtue*. Reprint Notre Dame, IN : University of Notre Dame Press (2007).
- Mauss, M. (1924). Essai sur le don. Forme et raison de l'échange dans les sociétés archaïques. *L'Année Sociologique*, 1. Translated as *The Gift: The Form and Reason for Exchange in Archaic Societies*, New York and London: W. W. Norton (1990).
- Mercier Ythier, J. (2010). Regular distributive efficiency and the distributive liberal social contract. *Journal of Public Economic Theory* 12: 943-78.
- Mercier Ythier (2011). Optimal redistribution in the distributive liberal social contract. Chap. 13 in M. Fleurbaey, M. Salles and J.A. Weymark eds., *Social Ethics and Normative Economics. Essays in honor of Serge-Christophe Kolm* (Berlin and Heidelberg: Springer-Verlag) 303-26.
- Mercier Ythier (2013). Virtue ethics and economics. Chap. 42 in L. Bruni and S. Zamagni eds., *Handbook on the Economics of Philanthropy, Reciprocity and Social Enterprise* (London: Edward Elgar) 430-40.
- Noddings, N. (1984). *Caring: A Feminine Approach to Ethics and Moral Education*. Berkeley: University of California Press.
- Nozick, R. (1974). *Anarchy, State and Utopia*. New York: Basic Books.
- Polanyi, K. (1944). *The Great Transformation*. Boston: Beacon Press.
- Posner, E. A. (1981). *The Economics of Justice*. Cambridge, M.A.: Harvard University Press.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge: Harvard University Press.
- Ricoeur, P. (1990a). *Soi-même comme un Autre*. Paris : Seuil. Translated as *Oneself as Another*, Chicago: The University of Chicago Press (1992).
- Ricoeur, P. (1990b). Ethique et Morale. In *Lectures*, t. 1, Paris: Seuil (1999).
- Ricoeur, P. (1990c). *Amour et Justice*. In *Amour et Justice*, Paris : Points (2008).
- Sacco, P. L., P. Vanin and S. Zamagni (2006). The economics of human relationships. Chap. 9 in S.-C. Kolm and J. Mercier Ythier eds., *Handbook of the Economics of Giving, Altruism and Reciprocity* (Amsterdam : North-Holland) 695-730.
- Savage, L. (1954). *The Foundations of Statistics*. New York: Wiley.
- Sen, A. (2009). *The Idea of Justice*. London: Penguin Books.
- Slote, M. (2001). *Morals from Motives*. Oxford: Oxford University Press.
- Smith, A. (1759). *The Theory of Moral Sentiments*. Reprint New York: Prometheus Books (2000).
- Thomas Aquinas (1981). *Summa Theologica*. Notre Dame, IN : Christian Classics.
- Vickrey, W. (1945). Measuring marginal utility by reactions to risk. *Econometrica* 13: 319-33.
- Von Neumann, J., and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton, N.J. : Princeton University Press.
- Walzer (1983). *Spheres of Justice: A Defense of Pluralism and Equality*. New York: Basic Books.
- Weymark, J.A. (1991). A reconsideration of the Harsanyi-Sen debate on utilitarianism. In J. Elster and J.E. Roemer eds, *Interpersonal Comparisons of Well-Being* (Cambridge: Cambridge University Press).