Justice as fairness, utilitarianism, and mixed conceptions

It would be hard to overstate the philosophical significance of John Rawls’s *TJ*. It articulates and defends an egalitarian conception of liberalism and distributive justice that consists of two principles of justice: a principle of equal basic liberties and a principle that distributes social and economic goods and opportunities so as to be to the greatest benefit of the least advantaged. Rawls defends this liberal egalitarian conception of justice primarily as an alternative to utilitarianism. He situates his defense of this liberal egalitarian conception of justice within the social contract tradition by arguing that his principles of distributive justice would be preferred to utilitarianism and other

Thanks to David Estlund, Tim Hinton, Theron Pummer, and Paul Weithman for comments on an earlier draft. I learned much of what I know about the appeal and resources of justice as fairness while I was at MIT in the late 1980s and early 1990s, especially from interactions with John Rawls, Tim Scanlon, Derek Parfit, and my (then) colleague Joshua Cohen. This essay develops ideas in lectures that I gave on justice as fairness during that period, which Tim Hinton attended and discussed with me.

1 Unless otherwise noted, all references to Rawls will be to *TJ*. The second edition (TJR) was published in 1999. Because the section numbering is the same in the two editions, I will refer to *TJ* using section numbers as much as possible. Rawls offers a restatement and reinterpretation of justice as fairness in *JFR*. Though *JFR* corrects *TJ* in places, its main point is to reinterpret justice as fairness as part of a specifically political conception of liberalism that aims to justify liberal essentials by identifying an overlapping consensus among different comprehensive religious, moral, and political commitments. This political reinterpretation of justice as fairness seems largely orthogonal to my main concern with the contractual argument for justice as fairness in *TJ*. I will take notice of *JFR* only insofar as it seems to bear directly on the reasoning within the contractual argument.

2 Rawls also defends justice as fairness against libertarian conceptions of justice. The libertarian conception gives moral significance, as justice as fairness does not, to the morally arbitrary effects of the natural and social lotteries – the distribution of natural talents and of social class and advantage (§§12–13). The argument against libertarian conceptions is an important part of the case for justice as fairness. Nonetheless, Rawls regards utilitarianism as the main rival to justice as fairness, both inside and outside the contractual argument, and the argument against libertarianism is largely a contract-independent argument. So in reconstructing and assessing Rawls’s contractual argument for justice as fairness, I will focus on the contrast with utilitarianism, largely ignoring his interesting and important discussion of libertarianism.
rivals by parties to a social contract in which they were represented fairly, that is, as free and equal moral persons. This explains why Rawls calls his conception *justice as fairness* and the importance of the hypothetical *original position* as a way of modeling a fair initial position from which principles of justice might be selected. Rawls thinks not only that justice as fairness would be preferred to utilitarianism in a fair social contract but also that it provides a better reconstruction than utilitarianism of our considered views about individual rights and justice. Though Rawls’s primary focus is on the justice of the basic structure of society, his critique of utilitarianism, his contractualist methodology, and his defense of equal basic rights have had much wider philosophical influence, extending to a variety of issues in ethical theory and normative ethics. In this way, the publication of *TJ* transformed and reinvigorated ethics as well as political philosophy.

Any assessment of justice as fairness must address the adequacy of Rawls’s contractual argument for his two principles of justice and against utilitarian rivals. In this context it is worth noting that Rawls has two kinds of ambition. On the one hand, he has the substantial but comparatively modest ambition to defend a more egalitarian alternative to utilitarianism. He would succeed in this ambition insofar as parties in the original position would indeed prefer his two principles of justice to traditional utilitarian rivals. On the other hand, Rawls also has the more ambitious aim of showing his two principles of justice to be uniquely plausible, that is, to be superior to all reasonable alternatives. He would succeed in this ambition insofar as parties in the original position would prefer his two principles of justice to any plausible rival. Basically, the more ambitious claim consists in claiming that Rawls’s two principles are superior to a greater number of rivals. Of special interest here is the possibility of rivals to Rawls’s two principles that differ from traditional forms of utilitarianism (whether classical or average) but nonetheless combine elements of utilitarianism with elements of liberal egalitarianism. Rawls calls these hybrids *mixed conceptions* (*TJ*, §21).

So one question is whether Rawls’s contractual argument succeeds against mixed conceptions, as well as against traditional utilitarian conceptions. Of course, the more modest claim that justice as fairness is superior to traditional forms of utilitarianism is a very important claim, whether or not Rawls can support his stronger claim that justice as fairness is uniquely plausible among reasonable alternatives. Moreover, the more modest claim would be positively relevant to establishing the stronger claim. But it is worth noticing that Rawls might succeed in the more modest aim without succeeding in the more ambitious claim if there are other rivals that justice as fairness does not
defeat. Rawls himself distinguishes between these two ambitions in the second edition of *TJ*, where he reasserts both ambitions but expresses greater confidence in the more modest comparative claim than in the more ambitious one (*TJR*, p. xiv). It is worth distinguishing these modest and ambitious claims not just to display logical possibilities. In fact, Rawls’s ambitious claim turns out to be much harder to justify than the more modest one, and his main contractual arguments for justice as fairness fail to demonstrate that it is uniquely reasonable, whether or not they succeed against more traditional utilitarian rivals.

In section 1.1 I will provide a brief reconstruction of the main elements of justice as fairness. In section 1.2 I will contrast Rawls’s general and special conceptions and ask whether something like the general conception isn’t a more plausible conception than the special conception, even when we restrict our attention to the circumstances that Rawls thinks justify adoption of the special conception. Then in section 1.3 I will reconstruct and assess Rawls’s specifically contractual arguments for his special conception of justice, focusing on the contrast between Rawls’s conception and traditional utilitarian principles. In section 1.4 I will focus on the adequacy of Rawls’s argument against a larger range of alternatives, including mixed and other conceptions. I will argue that even if Rawls can defend his more modest claim that justice as fairness is superior to traditional utilitarianism, his more ambitious claim that it is superior to mixed conceptions is problematic.

### 1.1 Justice as fairness: the two principles

One of the most distinctive features of justice as fairness is that Rawls develops an egalitarian conception of social justice that he defends by appeal to a hypothetical social contract, which he claims fits within the social contract tradition of Locke, Rousseau, and Kant. The contractual argument represents the basic structure of society and its provision of liberties, opportunities, and rights as just insofar as it satisfies principles that would be chosen in fair initial conditions in which contracting parties were represented as free and equal moral persons. It is essential to this sort of social contract that the contract be both *hypothetical* and *moralized*. The contractual argument

---

3 Similarly, in Part III of *JFR* Rawls distinguishes two “fundamental comparisons,” which correspond to the moderate and ambitious claims respectively. Though he still thinks that justice as fairness is superior to both unrestricted average utilitarianism and restricted average utilitarianism, he is more confident about the first comparison than the second.
requires Rawls to specify (a) fair initial contractual circumstances and (b) the principles that would be chosen in these circumstances.

We might begin with a preliminary specification of the principles that Rawls thinks would be chosen in this hypothetical contract. We will discuss issues introduced by Rawls’s distinction between general and special conceptions of justice later (section 1.2 below). But his primary focus is on the circumstances of justice faced by societies in conditions of moderate scarcity in which there is sufficient economic development and security to make possible a decent minimum standard of living for all (§26). For these conditions, Rawls defends what he calls the special conception of justice, which consists of two main principles (§§11, 46):

1 Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all [equal basic liberties].

2 Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, [the difference principle] and (b) attached to offices and positions open to all under conditions of fair equality of opportunity [fair equality of opportunity].

The currency of distributive justice for utilitarians is utility or well-being. For others, it is resources, or capabilities, or income. For Rawls the currency of distributive justice is primary goods – maximally flexible assets that it is supposed to be rational to want whatever else one wants (§15). Rawls adopts primary goods as the currency of distributive justice so as to avoid traditional worries about the basis for interpersonal comparisons of utility and to make the public deployment of his conception of justice easier, involving less complex calculations. Rawls understands primary goods to include rights and liberties, opportunities and powers, income and wealth, and the social bases of self-respect (cf. JFR, §17).

Notice that the three component principles here regulate the distribution of different goods: the equal basic liberties principle regulates basic liberties; the fair equality of opportunity principle regulates opportunities; and the difference principle regulates other kinds of social goods, including income.

The equal liberties principle concerns specific basic liberties, such as freedom of expression, conscience, religion, and association, rather than liberty per se. Insofar as liberties are prior to other primary goods, the most extensive system of equal liberty would lead to a form of libertarianism that would not
permit restrictions on liberty to prevent harm or nuisance. Because this is not Rawls’s claim, he must be concerned with specific liberties that seem fundamental. The first principle concerns those liberties relevant to the two primary capacities of citizens in a well-ordered society: (a) their sense of justice and (b) their interest in pursuing a conception of the good. Presumably, this includes liberties of thought, expression, and association, and political liberties, such as the right to vote, to campaign, and to run for office (cf. *JFR*, §13).

Rawls’s considered statement of the second principle emerges from a comparison of different interpretations of the more general idea that there be equality of opportunity and that any inequalities be to the advantage of all (§§12–13). Here Rawls describes a natural progression through several conceptions motivated by the idea that the terms of social cooperation should not make some worse off than they would otherwise be as the result of factors over which they have no control. This principle gives rise to worries about the system of natural liberty that combines laissez-faire with the idea that careers should be open to talents. The moral arbitrariness of the social lottery should force us to recognize the ways in which the idea of equal opportunity should constrain the unfettered accumulation of wealth across generations. But the same principle should make us skeptical of the system of liberal equality that allows unearned natural talents to determine a person’s social and economic prospects. The correct interpretation of opportunity and equality, Rawls thinks, should lead us to treat the distribution of natural talents as a common asset and insist on a conception of equality of opportunity that aims for reasonable accommodation of differences in native endowment.

The difference principle does treat the distribution of natural talents as a common asset (cf. *JFR*, §21). Robert Nozick believes that this implies that the community has property rights in the talents and powers of individuals, so that the community could compel talented individuals to work on the community’s terms. But Nozick’s criticism seems to conflate the following distinct ideas:

1. I am entitled to possess my natural endowments, though they are unearned.
2. I am entitled to exercise my natural endowments (in acceptable ways).
3. I am entitled to benefit from the exercise of my natural endowments (assuming they are productively employed).
4. I am entitled to the benefits that accrue from my exercise of my natural endowments in an unregulated market.

---

Rawls does not deny (1)–(3), only (4). Another way to see this involves distinguishing between rights of self-ownership and rights of worldly ownership, which we get from the Lockean tradition that Nozick embraces. Rawls does not think that the moral arbitrariness of the natural lottery raises questions about the rights of self-ownership of the talented, as Nozick seems to assume. Rather, Rawls thinks that the arbitrariness of the natural lottery affects the rules by which the talented can generate property rights in the world from the productive employment of the talents in which they have rights of self-ownership.

Rawls identifies two kinds of priority in his special conception of justice. The first principle is lexically prior to the second, so that liberty can only be restricted for the sake of liberty and the other goods cannot be distributed in ways that upset the demand for equal basic liberties. Moreover, the second part of the second principle is lexically prior to the first part, so that other social goods, including income, cannot be distributed in ways that upset fair equality of opportunity. Elsewhere, Rawls mentions another principle, which he calls the fair value of political liberties (§36). His idea here is to distinguish between the existence of a political right or liberty and its worth or value, where its worth or value can be affected by one’s absolute or comparative level of resources. We might have equal rights to make private expenditures on political campaigns, but this right will be of very unequal value to the rich and the poor. Rawls thinks we have compelling reason to care about the value of these political liberties. He does not insist on maintaining the equal value of basic political liberties, only on prohibiting unfair inequalities in their value. Rawls seems to think that ensuring the fair value of basic political liberties should be part of the priority of the first principle. Notice that Rawls refers to this principle as the fair value of political liberties. Since it is not clear why the concern with the worth of liberty should be restricted to political liberties, I will assume that it should be extended to all of the basic liberties.

If we were to display the components of Rawls’s special conception in a way that distinguishes different claims and reflects their priority relations, we might do it this way:

1 equal basic liberties;
2 fair value of basic liberties;
3 fair equality of opportunity;
4 the difference principle.

Rawls offers two very different kinds of defense of this conception of justice. He takes utilitarianism to be the chief rival to his more egalitarian conception
of justice. Some of Rawls’s arguments are arguments of overall comparative plausibility, independent of the contract argument. Some of these contract-independent arguments are quite powerful and have proved influential. But Rawls’s most distinctive arguments are contractual arguments that his principles of justice would be preferred to utilitarian rivals in the right sort of hypothetical moralized contract. Before focusing on these contractual arguments, it might be worth identifying some of the contract-independent arguments.

Utilitarianism, Rawls thinks, extends the decision procedure appropriate for intrapersonal contexts to interpersonal contexts. In doing so, it treats different lives as if they were parts of a single life, in which it makes sense to make individual sacrifices for the good of the whole. But whereas intrapersonal aggregation and sacrifice are acceptable, interpersonal aggregation and sacrifice fail to take seriously the separateness of persons. This is why utilitarianism has trouble recognizing individual rights, which trump utilitarian reasoning. In contrast with utilitarianism, justice as fairness places a limit on the sacrifices that some might have to make in order to improve the position of others and so recognizes rights as inviolable (§§5–6).

This is a complex and striking argument for Rawls’s liberal egalitarianism that deserves and has received extensive discussion elsewhere. In describing this as a contract-independent argument, I don’t mean to imply that it couldn’t play some role within his contractual argument. For instance, the critique of utilitarianism could enter into the contractual argument for Rawls’s two principles insofar as it might bear on the comparative strains of commitment and stability of rival schemes that contractors must assess (§§29, 76). My point is only that this critique of utilitarianism does not depend for its force on the contractual argument; it affects the overall comparative plausibility of Rawls’s conception and utilitarian conceptions.

This brings us to Rawls’s contractual argument for his liberal egalitarian conception of justice. The general form of the contractual argument is well known. Because the contract proceeds from a fair initial situation in which contractors are represented as free and equal moral persons, the bargaining situation is hypothetical and shaped by considerations of fairness. In particular, Rawls appeals to an original position that excludes factors that seem arbitrary from the moral point of view or that might be used to select principles favoring one group at the expense of another (§20). Parties in the

---

6 See e.g. Brink, “The Separateness.”
original position are placed behind a veil of ignorance, which deprives them of information about their identities and attributes, including their identity, gender, race and ethnicity, social class, level of social capital, generation, natural talents, attitude toward risk, and conception of the good ($24).

The deficits that the contracting parties operate under may make one wonder if they have the basis for preferring any principles and outcomes to others. However, the parties do have positive characteristics. They are mutually disinterested, but concerned to advance their own prospects; they use primary goods, which are maximally flexible social assets, to measure their prospects; they evaluate representative social positions, rather than individual lives; they display instrumental, means–ends, rationality; and they have general social knowledge about the forms of social, political, and economic organization ($§11, 24, 25).

The negative and positive attributes of the contracting parties combine to yield the following sort of choice situation:

Knowing all of the possible representative social positions that she might occupy in all possible societies at all possible points in time, a contractor should choose principles of justice that set terms of social cooperation that will best promote her stock of primary goods.

It is important to recognize that the original position is not supposed to be a description of how people actually are. It is an analytical device in which the characteristics of parties, considered together, represent a fair contractual situation such that any principles chosen in those circumstances to regulate the basic structure of society would be just. The parties in the original position do not represent Rawls’s picture of human nature; their characteristics are not intended to be descriptive individually or collectively. Together, the attributes of the contracting parties are supposed to model impartial and fair choice.

The opacity of the veil of ignorance appears to imply that it makes no difference to the contractual choice who the parties are. All parties, qua parties to the contract, would appear to be in the same position. This is why Rawls claims that the choice in the original position can be represented as a problem in individual decision theory under special circumstances, rather than as a contract or bargain among several parties with potentially conflicting interests ($§4, 20, 24).

Though the parties are not making a moral choice, they realize the point of their exercise and assess the consequences of alternative principles by considering their application in a well-ordered society ($§1, 69). In a well-ordered society, citizens have a sense of justice and a higher-order interest in pursuing
their conception of the good, whatever it turns out to be, once the veil of ignorance is lifted. Moreover, principles must satisfy a publicity condition, according to which the principles are recognized publicly as establishing standards of just institutional design and citizens actually assess their institutions in private and public fora by measuring their conformity to these principles (§§1, 23). Publicity, Rawls believes, imposes constraints of stability and practicality (§24).

Rawls is committed to the claim that the contracting parties will prefer his special conception of justice to utilitarian rivals. This is because utilitarianism, Rawls thinks, does not provide enough assurance against bad luck and exploitation by others. By contrast, Rawls’s two principles insure against both bad luck and exploitation. Later, we will examine and assess the details of this contractual argument. But perhaps this gives us a clear enough general picture to raise some questions about the principles that Rawls thinks would be chosen in the original position.

1.2 Special and general conceptions

So far, we have accepted Rawls’s focus on the special conception, which applies in societies above a certain developmental threshold, at which point Rawls thinks individuals care more for marginal increases in basic liberties than for marginal increases in social and economic goods. Below this developmental threshold, Rawls endorses the general conception, which basically distributes all goods according to the difference principle (§§11, 26, 46):

All social primary goods – liberty and opportunity, income and wealth, and the bases of self-respect – are to be distributed equally unless an unequal distribution of any or all of these goods is to the advantage of the least favored. (TJ, p. 303)7

Notice that the special conception differs from the general conception in two ways. First, it treats the distribution of some goods as lexically prior to others. Second, the special conception applies different distributional principles to these different goods: it insists on an equal distribution for liberties, fair value for the worth of basic (political) liberties, and fair equality for opportunities, while requiring that other goods be distributed according to the difference

7 Notice that this statement of the general conception does not explicitly require maximizing the position of the worst off, as the difference principle does.
principle. These two points of disagreement are independent of each other and should be assessed separately.

Rawls’s rationale for the special conception is based on the priority that he imagines contractors would assign to basic liberties once a certain threshold of material well-being was assured. But this claim about the priority of liberty supports at most the first difference between special and general conceptions. It does not support different distributional principles for the different goods. In particular, Rawls thinks that it is irrational for contracting parties to prefer equality to the difference principle when considering other primary goods (*TJ*, p. 151). But, if so, why not extend this reasoning to all the primary goods, while respecting the priorities that Rawls sees among them? No matter which primary goods we are talking about, it is not clear why contracting parties should prefer an equal distribution of those goods if there is an unequal distribution that makes everyone better off.

Recall that we said that when we factor in Rawls’s priority rules, we can see that the two principles that constitute the special conception are equivalent to the following ordered quartet:

1 equal basic liberties;
2 fair value of basic liberties;
3 fair equality of opportunity;
4 the difference principle.

Perhaps we should give liberties lexical priority over opportunities and other goods but distribute each by the difference principle. This would yield something like the following conception, which we might call the *extraspecial conception*:

1 difference principle for basic liberties;
2 difference principle for the value of basic liberties;
3 difference principle for opportunities;
4 difference principle for all other primary goods.

If Rawls has good reason to prefer the special conception to the general conception for circumstances of moderate scarcity, then perhaps he has good reason to prefer the extraspecial conception to the special conception.

Like the special conception, the extraspecial conception claims that in circumstances of moderate scarcity there should be strict priorities among different kinds of primary goods. But should we accept that claim? If there is a preference, above a certain threshold, for liberties over opportunities or other primary goods, surely it is not a lexical preference. Even if I wouldn’t trade
increases in basic liberties one-for-one with increases in other primary goods, surely I might prefer large increases in other primary goods to small increases in basic liberties? I might prefer to live much more comfortably, even if this would be purchased at a very small cost in my personal or political liberty. But if we reject both Rawls’s egalitarianism about liberties and opportunities (thus accepting the extraspecial conception in preference to the special conception) and his lexical priority rules, the result is something closer to the general conception, even for societies above the material threshold.

This is significant in itself inasmuch as we want to be clear about which principles Rawls should take his arguments to support. It is also potentially significant in assessing his more ambitious claim to defend his principles in preference to mixed conceptions, because the general conception is relevantly like a form of consequentialism in which priority is given to the worse off.

1.3 The contractual argument for the special conception

Why does Rawls think that the contracting parties would choose his special conception of justice in relation to utilitarianism and other alternatives? He presents the details of the contractual argument against traditional utilitarianism and in favor of the special conception in Chapter 3 (esp. §§26–9; cf. JFR, §§27–40). But he is committed to thinking that the parties have reason to prefer the special conception not just to utilitarianism but to mixed conceptions as well (§49).

One obstacle to comparing justice as fairness with utilitarianism and other rivals in the original position is that some of these rivals employ different metrics for assessing distributions from Rawls’s preferred metric of primary goods. Utilitarians assess distributions of utility or welfare, which would themselves be measured differently according to different conceptions of well-being. Other theories assess distributions of resources, opportunities, and functioning. We might focus on the difference between primary goods and utility. The principle of diminishing marginal utility claims that as a person increases her consumption of a given resource the marginal utility of additional units of that resource decreases. That principle ensures that distributions of primary goods and other resources do not in general reflect distributions of utility. This makes certain kinds of comparison between utilitarianism and Rawls’s principle difficult. But we can try to finesse this problem by focusing on distributions of primary goods in which we stipulate that these figures also represent equivalent differences in utility.
Within the contractual argument, Rawls takes utilitarianism to be the main rival to justice as fairness. We must distinguish between classical utilitarianism, which tells agents to maximize total happiness, and average utilitarianism, which tells agents to maximize average or per capita happiness.

Assume that D1 and D2 are two different possible distributions (see Table 1.1) and that A–D represent representative social positions, each of which has the same number of members (and D has no members in D1). Whereas classical utilitarianism endorses D2, because it has the highest total utility, average utilitarianism endorses D1, because it has the highest average utility. The contracting parties would prefer average to classical utilitarianism, because it maximizes their expected prospects (*TJ*, pp. 163–4).

Intuitively, Rawls’s argument against average utilitarianism and in favor of justice as fairness is that choosing the latter represents a more reasonable attitude to adopt toward risk in one’s life prospects. But this simplifies a more complex argument.

Rawls claims that for conditions of moderate scarcity the parties would prefer his two principles (the special conception) over both average utilitarianism and mixed conceptions. In arguing against mixed conceptions, he assumes that we have already satisfied the prior principles (equal basic liberties, fair value of political liberties, and fair equality of opportunity) and argues that the parties would then choose the difference principle over a mixed conception employing average utilitarianism in the role that justice as fairness assigns to the difference principle. If justice as fairness is superior to such a mixed utilitarian conception, then a fortiori it will be superior to unrestricted average utilitarianism.

Rawls makes this case by contrasting two decision rules: the maximin rule that instructs one to make the worst outcome as good as possible and the rule of maximizing expected utility. These two decision rules will evaluate possible distributions differently.

Assume again that the matrix in Table 1.2 represents possible distributions of both primary goods and utility (circumstances in which the two coincide) among three representative groups. The question is whether parties should

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>D2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
maximize expected utility, gambling that they will be one of the better off in D1 or D3, or whether they should minimize risk and ensure the best worst-case scenario by choosing D2. Maximin supports adoption of the difference principle, and the principle of maximizing expected utility supports adoption of average utilitarianism.

Rawls says there are three conditions that favor the use of maximin, though it is not entirely clear whether they are supposed to be individually or jointly sufficient for this purpose (§26; cf. JFR, §28):

1. ignorance of the probabilities of the various outcomes;
2. a conception of the good such that one cares little, if anything, for what one might gain above a certain minimum;
3. alternative decision rules have possible outcomes that would be unacceptable.

It seems crucial to Rawls’s argument that these three conditions are satisfied or approximated in the original position. So, in assessing this argument, we should distinguish two kinds of issue: whether these conditions, individually or collectively, favor maximin, and whether these conditions are satisfied or approximated in the original position in an appropriate way.

Consider (1). Standard decision theory distinguishes decisions under three different epistemic conditions:

A. **Certainty** in which the probability of the outcome is 1 on a scale of 0–1.
B. **Risk** in which the probability of the outcome is fixed but less than 1 or in which there is a probability space (e.g. 0.5–0.6).
C. **Uncertainty** or **ignorance** in which there is no determinate information about the probability of the outcome.

Standard decision theory identifies rational decision-making with maximizing expected utility. This is straightforward in contexts of certainty – prefer more utility to less. It is also the conventional view of rational decision-making in contexts of risk. For instance, one should prefer 1/3 chance of 100 utiles to a 1/2 chance of 50 utiles to a 1/1 chance of 20 utiles. However, it is usually

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1</td>
<td>10</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>D2</td>
<td>12</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>D3</td>
<td>5</td>
<td>15</td>
<td>26</td>
</tr>
</tbody>
</table>
thought that rational decision-making is impossible or undefined for uncertainty or ignorance. Under such circumstances, the principle of maximizing expected utility has no application.

Rawls recognizes that maximin is not normally a rational decision rule (TJ, p. 153). If we adopted it generally, we would never take risks, no matter how low their probability or how great the possible payoff. He does not advocate using maximin in conditions of certainty or risk. But he does seem to think that maximin might sometimes be rational under uncertainty.

This is part of what makes it important that the choice in original position is supposed to be under uncertainty. John Harsanyi thought that contracting parties should assume that the various possible outcomes were equiprobable. If we make this assumption about the original position, then the choice in original position becomes a decision under risk. So understood, Harsanyi argued, parties should employ the principle of maximizing expected utility and should therefore adopt average utilitarianism. However, Rawls understands the choice in the original position as one involving uncertainty, rather than risk. The contractors are not supposed to have knowledge or even beliefs about the comparative likelihood that they will occupy various representative social positions or historical circumstances. Rawls is right insofar as he claims that the argument for average utilitarianism is undermined if the choice in the original position is one under uncertainty, rather than risk.

But that leaves two potential gaps in Rawls's argument. First, even if it is not rational to employ the principle of maximizing expected utility under conditions of uncertainty, it does not follow that maximin is uniquely rational. Perhaps rationality is just undefined in these circumstances. Certainly, further argument is required to show that maximin is uniquely rational under uncertainty. Second, there is the question of whether the original position should be modeled in terms of uncertainty or risk, such as equiprobability. We might claim that insofar as our task is to model what principles free and equal persons would choose to govern the basic structure of their society we should represent the various possible outcomes and our roles in them as equally probable. Similarly, if we thought that the point of the original position was to represent members of society impartially, we might argue that this would be best achieved by asking them to treat outcomes as equally likely, whoever they turned out to be when the veil of ignorance was lifted. To block these utilitarian arguments, Rawls needs to provide a more

---

8 Harsanyi, “Morality.”
satisfactory rationale for claiming that fairness requires uncertainty, rather than equiprobability. Perhaps he could make use of his idea that fairness requires excluding from the original position information that could be used to form coalitions that might disadvantage some at the expense of others. But the flip side of excluding information that can be used to advantage some at the expense of others is excluding information that might prevent some from holding others hostage to their own benefit. Again, I think, the demands of fairness here are unclear.

Now consider condition (2), which stipulates that one cares little, if anything, for what one might gain above a certain minimum. As it stands, (2) does not help the case for maximin. The conception of the good must exhibit a marginal utility threshold above which one cares very little or nothing for further increments in primary goods, and – importantly – this utility threshold must be at or below the minimum that maximin would secure. If the utility threshold is above the minimum that maximin would secure, then the utility threshold won’t favor maximin over maximizing expected utility. Moreover, it is not enough that above this threshold one cares less for further increments in primary goods; above this threshold, one must care absolutely nothing for further increments of primary goods. So the assumption that Rawls really needs is something like this:

2’ A conception of the good such that one cares nothing at all for increments above the highest minimum that one might receive.

In questioning Rawls’s priority rules in the special conception, we have already suggested that this claim is extreme.

This poses two problems for Rawls. First, it is no surprise that this conception of the good would support maximin. If the original position were to build this assumption in, the contractual argument would be circular – we would be deriving maximin from a version of the original position that presupposes maximin. Second, Rawls stipulates that contracting parties do not know their conception of the good (TJ, pp. 18 and 137). A fortiori they do not know that they have this peculiar conception of the good.

The first two rationales for maximin are not very compelling. So let’s consider (3)’s claim that alternative decision rules have possible outcomes that are unacceptable. This is perhaps Rawls’s strongest rationale for use of maximin in the original position (JFR, §29). His idea seems to be that the stakes here are extremely high, because they concern one’s life prospects. We accept certain risks and take gambles where the stakes are smaller and where we can absorb losses fairly easily. But as the stakes become greater,
Rawls reasons, it is harder to accept unnecessary risk. If one takes an unnecessary risk when the stakes concern one’s life prospects then losses become unacceptable. Rawls seems to think that it will seem even more unacceptable to lose a gamble when there was a less risky alternative available. For example, if one gambles that one will be a master in the course of embracing the permissibility of slavery, when the option of prohibiting slavery was available, then the prospect of becoming a slave will be unacceptable or unbearable. So the idea is that in the selection of life prospects the cost of losing a gamble that one ends up as one of the Haves and turns out instead to be a Have Not is not just bad but unacceptable, especially when one could have employed maximin and ensured that the worst-case scenario was least bad. Since the contractual choice in the original position is precisely a selection of principles governing life prospects, this favors maximin, rather than maximizing expected average utility.

Though Rawls is right to call attention to the high stakes involved in the choice in the original position, this rationale for maximin is problematic. For one thing, the high stakes cut two ways. Not only is there more to lose in a gamble, but there is also more to gain. Moreover, maximin appears not to be a reasonable attitude toward risk, even when some losses are catastrophic. Suppose that I suffer from a condition that involves significant, persistent pain but that my life is otherwise worth living and valuable. Suppose, moreover, that I can have surgery performed to relieve my pain that has an extremely high success rate. Nonetheless, surgery always carries a small risk of catastrophic outcomes of various kinds. It doesn’t seem reasonable to forgo the kind of surgical treatment that offers a very good chance of significant relief on the ground that some possible outcomes of the procedure are catastrophic. The risk of catastrophic outcomes that are possible but unlikely is not always or even usually unacceptable, certainly not ex ante.

Finally, notice that Rawls writes as if there are only two options—maximizing expected average utility in a way that tolerates considerable risk if the gains are great enough and maximin. But, of course, maximin forbids nearly all risk-taking. So it rules out not only risking catastrophic outcomes but also outcomes that are risky but where the risk is limited in various ways. It is hard to see how that degree of aversion to risk could be the product of choice in the original position unless the parties are stipulated to be highly averse to risk. But that would beg the question against less extremely risk-averse principles and would, in any case, be inconsistent with Rawls’s insistence that contracting parties should be ignorant of their individual attitudes to risk (TJ, p. 137).
1.4 The comparison with mixed and other conceptions

This last point is relevant to assessing Rawls’s more ambitious claim to provide a contractual justification of his two principles (the special conception) not just in relation to unrestricted average utilitarianism but also in relation to other possible conceptions of justice, including mixed conceptions. Recall that in the special conception of justice as fairness the difference principle is posterior to three prior principles:

1. equal basic liberties;
2. fair value of basic liberties;
3. fair equality of opportunity;
4. the difference principle.

According to the special conception, the difference principle distributes other primary goods only after and in ways constrained by the equal basic liberties, the fair value of basic liberties, and the fair equality of opportunity.

Mixed conceptions are conceptions of justice that differ from Rawls’s special conception only with respect to (4). In other words, mixed conceptions must also respect these prior principles, (1)–(3). But that means we need to distinguish two different sorts of utilitarian rivals to Rawls’s own special conception: (a) unrestricted average utilitarianism, which does not recognize any prior principles that constrain the maximization of expected average utility, and (b) a mixed conception, which we might call mixed utilitarianism, which enjoins us to maximize expected average utility in ways constrained by the three prior principles. For comparison, mixed utilitarianism would embrace this ordered quartet:

1. equal basic liberties;
2. fair value of basic liberties;
3. fair equality of opportunity;
4. average utilitarianism.

Of course, mixed utilitarianism is only one kind of mixed conception.

These distinctions are important, because Rawls’s clearest contractual argument against utilitarianism works at most against unrestricted average utilitarianism. This is his argument that in the original position contractors should be loath to take the risks involved in maximizing expected average utility, because possible outcomes can be completely unacceptable. The example he gives is that it is conceivable that a slave society might maximize average utility and that we might end up as slaves in such a society once the
veil of ignorance was lifted (TJ, pp. 156 and 167–8). Though the principle of diminishing marginal utility makes it very unlikely that a system of slavery would maximize average utility, perhaps the conceptual possibility that it might is reason enough to condemn average utilitarianism. But, of course, slavery would violate the equal basic liberties principle. Because slavery would violate a prior principle, mixed utilitarianism would necessarily prohibit these social arrangements. More generally, the prior principles constrain permissible social and economic inequalities in significant ways. Inequalities in primary goods that would affect the equal provision of basic liberties, their fair value, or fair equality of opportunity are ruled out by mixed utilitarianism. Exactly how big a substantive constraint the prior principles impose is an interesting and complex issue about which there can be reasonable disagreement. But if we think about the way in which significant differences in wealth can compromise fair equality of opportunity at a time or across generations, it seems likely that these prior principles will significantly constrain permissible social and economic inequality. Prior principles will likely both set a threshold below which people cannot fall and significantly constrain possible inequalities among people above that threshold. But then prior principles limit the amount and extent of gambling that mixed utilitarianism might countenance. If so, it is much harder to make out that contracting parties should prefer Rawls’s special conception to mixed utilitarianism, because it is hard to see mixed utilitarianism tolerating outcomes that are genuinely unacceptable.

Rawls came to recognize that the argument against unrestricted average utilitarianism does not work against mixed utilitarianism, which ensures against possible outcomes that are unacceptable (JFR, §§27–34). He would later argue against mixed utilitarianism and in favor of his special conception by appeal to the social bases of self-respect, reciprocity, publicity, and stability (JFR, §§34–8). 9 But these arguments are largely independent of the contractual argument, and even Rawls recognized that they are not decisive (JFR, §40). It is not clear why the social bases of self-respect and reciprocity are not addressed by the prior principles that demand equal basic liberties, the fair value of basic liberties, and fair equality of opportunity. Indeed, one might wonder whether guaranteed distributional minima beyond those contained in the prior principles are fully compatible with a culture of mutual respect and reciprocity. It is hard to see how mixed utilitarianism undermines publicity

9 Also see Cohen, “Democratic Equality.”
any more than the special conception. While it is true that mixed utilitarianism may allow greater dispersion of income than the special conception, it severely limits permissible dispersion. So it is not obvious that mixed utilitarianism generates troublesome strains of commitment. Moreover, we must remember that precisely because it further constrains income dispersion, beyond that required by the prior principles, the special conception may give rise to its own strains of commitment among those whose life prospects would have been better had the basic structure not been required (by the difference principle) to provide guaranteed minima above those required by the prior principles. More generally, Rawls faces a challenge here explaining why these values (self-respect, reciprocity, publicity, and stability) require more constraints on economic dispersion than the substantial constraints guaranteed by mixed conceptions but fewer constraints than those imposed by an equal distribution. It’s hard to see how so much could hang on threading that needle.

We might compare mixed utilitarianism with other conceptions that are structurally similar. One such mixed conception is *sufficientarianism*. Sufficientarianism focuses not on securing equality, as such, or giving priority to the worst off, as such, but rather on ensuring that people fare *well enough*. There are different ways of articulating sufficientarianism, but the central idea is to recognize a *moral asymmetry, relative to some threshold level of welfare* such that all else being equal the urgency of benefits below the threshold is greater than the urgency of benefits above the threshold. Different sufficientarian conceptions will specify the threshold differently and will have to weight and combine independent variables, such as a beneficiary’s distance from the threshold, the size of the benefit to be conferred, and the number of similarly situated people that can be benefitted. However these details are spelled out, sufficientarian views will disagree with Rawls’s special conception and, in particular, the difference principle. All else being equal, they will prioritize benefits to the worst off but only up until the point of sufficiency. On some sufficientarian views, there might be no special concern for the worst off above the sufficiency threshold. Perhaps above that threshold, we should maximize expected average utility. Provided that the threshold established by this form of sufficientarianism is sufficiently close to the baseline established by the prior principles, this form of sufficientarianism might be quite similar to mixed utilitarianism.11

---

  For some skepticism, see Arneson, “Distributive Justice” and Casal, “Why Sufficiency.”
11 Cf. Waldron, “John Rawls.”
This and other forms of sufficienarianism provide a response to Rawls’s appeal to the unacceptability of possible outcomes in his argument against average utilitarianism and in favor of the special conception. Even if that was a good argument against unrestricted average utilitarianism, it would not support the special conception, because there are other ways of ensuring against worst-case outcomes that don’t require committing to the difference principle. Moreover, because the prior principles insure against many intolerable outcomes, there are also mixed conceptions, such as mixed utilitarianism, that provide viable contractualist alternatives to the special conception.

1.5 Concluding remarks

We should distinguish between Rawls’s justificatory appeal to a hypothetical moralized contract to derive principles of justice and his claim that such a contract uniquely favors his form of liberal egalitarianism, enshrined in the special conception of justice as fairness. There is reason to be skeptical of Rawls’s own principles. First, Rawlsian arguments might provide a better fit with his general conception of justice, rather than his special conception, even for the circumstances of moderate scarcity for which he designed the special conception. Second, it is not clear that it is reasonable for contracting parties to employ maximin, which raises questions about Rawls’s more modest aim of defending his special conception in preference to traditional utilitarianism, in particular, unrestricted average utilitarianism. Third, even if Rawls could defend this more modest claim, he could not defend the more ambitious claim that the special conception should be preferred to all plausible rivals, including various mixed conceptions. In mixed conceptions, prior principles set a decent social minimum below which no one can fall and limit social and economic inequalities above this minimum. This means that even if unrestricted average utilitarianism countenances possible outcomes that are unacceptable to the worst off, mixed utilitarianism does not. Similarly, most forms of sufficienarianism prioritize establishing a decent minimum for all, which means that Rawls’s best argument against unrestricted average utilitarianism is not a good argument against sufficienarianism.

This skepticism about Rawls’s contractual defense of justice as fairness against utilitarianism and mixed conceptions should not obscure the philosophical significance of justice as fairness. In particular, even if this skepticism is correct, Rawls’s defense of liberal egalitarianism could proceed in one
of two ways. One could eschew the contractual argument and rely on contract-independent arguments to criticize utilitarianism and defend his special conception. Alternatively, one could maintain the contractual argument and its appeal to the original position and explore more fully just which principles of justice that supports, paying greater attention to the merits of sufficientarianism and mixed conceptions.