

Against the Difference Principle

In *A Theory of Justice*, American moral and political philosopher John Rawls uses his notorious "justice as fairness" model to derive what he contends to be an essential principle for a just society: the difference principle. In order to determine the best guiding principles for a society, Rawls describes a hypothetical situation, known as the "original position," in which equal, rational agents are to agree on a proper conception of justice. To eliminate potential bias from the derivation of these principles, Rawls imagines agents of the original position to be subjected to a "veil of ignorance" such that they are ignorant of any facts of their personal situations. Furthermore, agents are also ignorant of any *probabilities* of their potential standing in society so as to eliminate any strategic calculations in their assessment of principles. According to Rawls, agents in the original position would agree to the well-known difference principle which holds that social and economic inequalities of a society are just only if they maximize the benefit of the least advantaged members. While I do not dispute the ability of Rawls' social contract model to determine justifiable guiding principles for a society, I do, however, disagree that the difference principle would be agreed upon in the original position.

Breakdown (Modus Ponens)

P1: If it is not evident that the difference principle would be agreed upon in the original position, we need not accept it as an essential part of justice.

P2: It is not evident that the difference principle would be agreed to in the original position.

C: We do not need to accept the difference principle as essential to justice.

Justification of Premise 1

Since Rawlsian contract theory contends that only those principles agreed to under the conditions posed by the original position should define justice, we need not accept the difference principle if its derivation in the original position is questionable. Put differently, while a critic of the difference principle may reject it and Rawls' contract theory entirely by appealing to a different distributive justice theory (i.e. utilitarianism), one sure-fire way to refute the difference principle is to show that it would not be rational to adopt it even in Rawls' contractual situation.

Justification of Premise 2

This section gives multiple explanations for why rational individuals in the original position are unlikely to agree to the difference principle.

If adopted, the difference principle has the potential to disincentivize the development of certain traits that are essential to a productive society. As aforementioned, a central component to the original position is a veil of ignorance such that participants are unaware of certain facts such as sex, wealth, and personal traits that could bias them in their support of certain

principles. For example, if a rational individual knew they were particularly productive and ambitious, they may favor principles that permit greater inequality, and if they knew they were white, they may favor a prejudicial society. However, the difference between these contingencies is that race is a permanent attribute, while an individual's values, work ethic, disposition, and so forth can change with time. Furthermore, the fluid nature of these particular attributes is presumably known to those in the original position as this is a general fact of human psychology. By subjecting individuals to the difference principle and thereby reducing the potential gain an individual receives from having certain traits associated with productivity (ambition, determination, etc.), individuals would be disincentivized to develop these traits that are objectively valuable. As such, even without knowing any facts about their own personal character, a rational participant of the original position may not consent to the difference principle based on the notion that it could reduce the probability of them acquiring these desirable attributes and values. Furthermore, with fewer individuals reaching their full productive potential, society as a whole, including the least advantaged, could be made worse-off demonstrating the inefficacy of the difference principle.

In his derivation of the difference principle, Rawls does not adequately account for the fact that social primary goods have a subjective value that varies from individual to individual. After establishing the premise that all humans are of equal moral worth, Rawls implicitly, and incorrectly extends this notion of equality to human judgements. Put differently, in his conjecture that the difference principle is the maximin solution to the problem of social justice, Rawls quantifies different potential outcomes of individuals in society solely in terms of social primary goods without considering that humans themselves value these things differently. To illustrate this point, imagine a society with a completely equal distribution of wealth and power. Furthermore, imagine that for one reason or another, this society's redistributive policies have maximized output such that Rawls would consider this a maximin solution to the problem of social justice and therefore a just society. While this setup may maximize the economic and social standing of the "least advantaged," there would still be considerable disparities amongst individuals in terms of their well-being. For example, an individual in this society whose happiness is highly correlated to the practice of an artform would be considerably better off than an individual whose material wealth and success is a vital aspect of their self-actualization. In effect, the difference principle may discriminate against those who weigh primary social goods more in their considerations for well-being. This rather abstract notion of "well-being," or an individual's happiness is a far more important consideration to economic and social standing as these things are only relevant insofar as they improve an individual's well-being. Rawls' maximin explanation for the difference principle is inadequate as different individuals place different values on social primary goods which is presumably known by agents in the original position. As such, individuals in the original position would account for the fact that they may be an individual who places greater weight on materialism and reject the difference principle on the basis that it could potentially interfere with their pursuit of happiness. In essence, Rawls' maximin derivation of the difference principle implies that human outcomes, values, and welfare are quantifiable when they are entirely subjective and unknowable but only to the self.

Objections

Objection 1: *By allowing room for some inequality, the difference principle does not reduce the incentives of individuals to work hard.*

Rebuttal: The value we place on anything, whether it be a trait, object, or experience is based on the return we receive from it. If an individual receives \$X as a result of her unrelenting ambition in a democratic society devoid of the difference principle, any diminishment in X as a result of implementation of the difference principle would result in the devaluation of her ambition. Though Rawls does not specify exact policies that would follow from the difference principle, it is likely that the difference principle would significantly reduce X now that the individual must not act solely in self-interest, but must benefit everyone in her actions.

Objection 2: *By allowing room for some inequality, the difference principle does not reduce the ability of success-oriented individuals to self-actualize*

Rebuttal: Returning to the aforementioned example of a completely equal society with redistributive policies that have effectively maximized production, we can also imagine that, by nature of these effective policies, any inequality of wealth will reduce the wealth of others thereby worsening the condition of the "least advantaged." Given that any inequality of wealth would take the society off the maximin frontier, under the difference principle no inequality of wealth would be allowed whatsoever. As such, we can imagine an individual whose sole goal and purpose in life is to reach a certain level of wealth that is unattainable under the difference principle. While one may contend that the scenario depicted in this counterexample is extreme and/or unlikely, one of the conditions of Rawls' original position (and maximin process) is that the probabilities of certain outcomes are unknown. Therefore, while a more reasonable example could make the same point, this example is no more illegitimate.

Objection 3: *While it is true that an individual's "happiness" or "well-being" is subjective and specific to a particular person, primary social goods grant individual's more time to partake in activities or obtain goods that make them happy. Therefore the maximin process does not have to consider the "well-beingness" of societal outcomes in the maximin process.*

Rebuttal: While it is true that to some degree everyone will be made better off given additional primary social goods (i.e. marginal utility of a dollar), this marginal utility differs from person to person and is also impossible to quantify. This difference of marginal utility can help us understand why some individuals work harder than others professionally; they value what they are working for more. Put differently, while it is true that primary social goods, particularly money, will make anyone better off, the subjective values different people have for these social primary goods is not taken into account by Rawls.