
Eric G. Roscoe, American University
JOHN RAWLS: PRINCIPLES THAT MAKE SENSE.

By Eric G. Roscoe

A Junior Independent Study Thesis
Submitted to the Department of Political Science
at The College of Wooster
December, 2005
In fulfillment of the requirement of Junior I.S. Thesis

Advisor: Mark Weaver
<table>
<thead>
<tr>
<th>CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
</tr>
<tr>
<td>The Original Position</td>
</tr>
<tr>
<td>The Two Principles of Justice</td>
</tr>
<tr>
<td>Critiques</td>
</tr>
<tr>
<td>Rawls’ Response</td>
</tr>
<tr>
<td>Conclusion</td>
</tr>
<tr>
<td>Works Cited</td>
</tr>
</tbody>
</table>
Introduction

John Rawls stands as one of the most prominent political theorists of the past three decades. His magnum opus, *A Theory of Justice*, was an attempt to find a balance between equality and liberty through the use of a social contract. There was great controversy caused by the nature of Rawls’ theory, a universal moral theory derived through the social contract tradition, and the controversy resulted in an encyclopedia of criticisms. Rawls even in the first published edition of his theory attempts to provide answers to these criticisms; he tangles his theory throughout with explaining the design, and then going back and rehashing in order to secure it from critiques.

In order to understand Rawls one needs to begin with an understanding of why Rawls has chosen to go the route of writing a comprehensive moral theory through the use of a social contract, something not really executed since John Locke. For Rawls there exist theories that he believes do not fully answer questions of justice for society, and he begins his analysis with critiques of the more prevalent forms, utilitarianism and intuitionism. Rawls critiques each theory on the basis that these other theories contain inherent flaws that prevent them from standing on their own merits. It is necessary to provide an overview of Rawls’ critiques of utilitarianism and intuitionism, because *A Theory of Justice* should be viewed as an answer to the flaws inherent in these moral theories.

Rawls critiques these theories based on the premise stated above that given a goal of maximizing certain goods these conceptions of justice will not fully maximize the goods necessary for a just society. They are two distinct ways of conceiving justice, and Rawls uses their definitions in determining a base for his own principles in order to prove the superiority of

---

1 In all of the critiques of moral theories that Rawls provides it is important to note that he is searching for something that is lacking, and he believes that his principles will fill in these areas of discord.
his theory. He defines utilitarianism, as a societal system in which justice is seen when the major institutions are arranged so as to provide the greatest amount of happiness for the greatest number of individuals involved. (Rawls 1999, 20) Rawls finds utilitarianism to be inadequate as a moral theory because, it turns on our judgments in regards to value. It enables individuals to make judgments in regards to situations without considering the rights of all individuals. (Rawls 1999, 22) Rawls is concerned that there is the possibility that a situation might arise where the sacrifice of certain liberties or rights would provide for a greater good to society.

Whenever a society sets out to maximize the sum of intrinsic value or the net balance of the satisfaction of interests, it is liable to find that the denial of liberty for some is justified in the name of this single end. (Rawls 1999, 185)

Utilitarianism can lead to a situation where incredible harm and displeasure even the torture of one person may result in the greatest good to society as whole. Therefore severe injustice committed against one individual can be justified on the grounds that it provides happiness to all other individuals. Clearly a conception that justifies the torture and suffering of a minority for the pleasure and happiness of the majority is not a viable moral theory.

In order to solve the discrepancy created in judgments of value, Rawls discusses how utilitarians are forced to adopt the concept of an impartial spectator. This one individual is charged with the task of organizing everyone’s desires into one “coherent system of desire.” (Rawls 1999, 24) It can best be described as a situation in which pains are intrasubjectively shared. For example, given two similar situations, one person may find being cut on the arm very painful and will experience great displeasure, while another person may not even notice the cut. Happiness, pleasure, pain, and joy are all value-laden emotions that rely on the person.
Utilitarians are therefore forced to posit an impartial spectator, someone who decides what is considered harmful and what is not, since there can never be hope for unanimity in values.

Since there is only one person that determines the best distribution of desires Rawls argues that, “Utilitarianism does not take seriously the distinctions between persons.” (Rawls 1999, 24) Rawls thus finds two discrepancies, which given the idea of maximizing certain concepts, cause the utilitarian conception of justice to fail as a thorough solution. First, rights and freedoms are of value, and a desire to protect all with a moral theory should be tantamount. So there is no maximization and protection of liberties, thus utilitarianism is rejected as a theory that permits grave injustice against some for the benefit of a majority. Secondly, the impartial spectator model needs to be used since there is no unanimity on values. An impartial spectator permits one individual to design a system of justice, this system does not allow for full participation of other persons, and creates a conception of justice limited by time, space, and social position.

Intuitionism creates a whole new set of problems for Rawls because it goes against the whole goal of moral theory. First, intuitionism consists of a “plurality of first principles,” which can conflict, and work in complete contradiction to each other. These principles therefore provide no basis for making moral decisions. Secondly, there are no rules, no procedure for weighing the importance of the principles. In Rawls’ words, “we are simply to strike a balance by intuition, by what seems to us most nearly right.” (Rawls 1999, 30)

The main fault Rawls finds is that intuitionism is a relativist theory, it creates these different balances of principles that apply to different individuals. Rawls is very Kantian in the sense that he searches for a moral individual, someone that exists without all the societal influences. (Hampton 1980, 336) This Kantian ideal allows Rawls to believe in the idea that
there are certain conditions necessary in order to obtain a moral theory or solution. However, he believes that there can be even deeper meaning than just universality derived from Kant’s notion of autonomy, which is addressed by Rawls in his analysis of individuals in the Original Position. (Rawls 1999, 221)

Rawls argues that there are no substantive criteria for guiding one in choosing the correct principles under intuitionism, and this leads different persons to arrive at different balances of principles. Since none of the principles are lexically ordered then intuitionism fails to answer the “priority problem.” (Rawls 1999, 34-36) The priority problem for Rawls is that principles need to be prioritized in order to provide guidance in making moral decisions. The weighing and ordering of principles helps to ensure a priority of liberty, and also guarantee that given strict adherence to the ordering of the principles we can obtain a certain equality of opportunity.

So, we can see two real problems with intuitionism that justify dismissing it as an incomplete conception of justice: 1.) First, there is no set of principles, which guide moral thought; rather there is a plurality of conflicting principles. Therefore, intuitionism fails by not providing universally acceptable principles for judging moral dilemmas. 2.) Secondly, there is no ordering of the principles thus creating a problem similar to that of utilitarianism, meaning that there is not necessarily any prioritizing of the right.

It is with the moral theories of utilitarianism, intuitionism, and others that Rawls derives a series of conceptions of justice, which he feels his principles will maximize in all areas thus picking up for deficiencies in other conceptions. Rawls believes we can obtain a universally applicable moral theory through the use of a social contract. The individuals entering the Original Position will ultimately choose principles that mesh with a particular conception in

---

2 Within an essay by Oliver A. Johnson, *The Kantian Interpretation*, Johnson critiques the legitimacy of Rawls’ Kantian design. He examines the Rawlsian use of Kant’s autonomy, the categorical imperative, and rationality, and concludes that Rawls fails in following the proper Kantian definitions of these terms. (Johnson 1974, 58)
which one maximizes certain moral principles. Principles that current popular ideologies are insufficiently equipped to deal with.

For Rawls these two moral theories are completely unacceptable given the nature of the Original Position shaped in *A Theory of Justice*. The critique of these two moral theories is where Rawls begins his conception of justice. He argues a solution to the problem of providing a universally applicable theory that provides answers to moral dilemmas. His theory is deeply layered around the initial premise that current theory is insufficient, and that the problems themselves can be answered by looking towards the old methods of contract theory. Exposing the layers and analyzing Rawls’ main arguments is where I begin my discussion.

I will present my analysis of Rawls in four parts: 1.) Starting with an analytical explanation of *A Theory of Justice*, I will examine the work in two parts: a. the Original Position, and b. the Principles of Justice. 2.) I will then provide several of the larger critiques in their entirety, and look at how accurately they handle the design of Rawls’ theory. At this point I will provide my own critique of Rawls as well, focusing in especially on the design of his Original Position. 3.) I then examine the response that Rawls formulates to his critics in *Justice as Fairness*. 4.) Finally, I conclude that the principles themselves can stand alone without the OP because of a separate grounding, and I question the neutral moral grounding of the Original Position.

I start with a simplistic breakdown of Rawls’ theory into two categories, chapter one The Original Position and chapter two The Principles of Justice. Throughout the first two chapters I will be attempting to extract an ordered argument from Rawls’ original theory. I believe that there is a flow to the theory that Rawls provides. If one follows his premises and believe that there are certain human attributes necessary for determining moral principles, then one can see
how and why his theory takes shape. In writing *A Theory of Justice* Rawls layers his analysis under rebuttals to different critiques. The constant rehashing of points used to develop the theory actually disguises certain flaws in the basic premises of the argument. My goal here is to pull out the basic premise, flaws and all, so that the critiques can be analyzed based on their relevance to the theory itself, and not the Rawlsian answers to critiques.

In chapter one I break the Original Position, which as the starting point for his social contract, into three distinct areas: 1.) First, I analyze the nature of individuals that enter the Original Position. The actors have what I categorize as a Rawlsian conception of Human Nature. The Rawlsian conception rests heavily on a Kantian notion of the self. He bases the original position on a unique conception of Emanuel Kant’s categorical imperative, and Hume’s idea of the rational and impartial sympathetic spectator. 2.) Second, I follow the individuals into the Original Position and analyze what the “conditions” or rules imposed on actors in the Original Position. 3.) Finally, I conclude by examining how all of this leads to the maxmin rule, which results in the two principles of Justice.

Within the Original Position Rawls states that we arrive at two principles of justice, the liberty and difference principles. These principles are the result of deliberation between the parties in the Original Position, and they represent universalizable moral positions. The two basic principles are a starting point from which a government that favors equality and justice can be derived. Rawls believes that with these principles the representative individuals could then enter into a constitutional convention, and deduce general liberties and rights that would exist in a truly just society.

I will analyze the definitions of the principles of justice in the second chapter, but it is important for an understanding of the first chapter to understand the basic concepts behind each
principle. Rawls introduces the principles first in his theory, however, I feel the inevitability of the principles being chosen under the Original Position is lost when he does this, and it appears as if he is just formulating a position to match his principles. I know Rawls believes that his Original Position is not designed to match the principles, but rather it is designed to find the most just principles, and it happens that his two principles are what are derived from the Original Position. I am presenting only a brief overview of the principles in an attempt to show that the Original Position provides the basis for determining moral principles and serves as a thought experiment for selecting principles of justice. Specifically the Original Position tries to answer the problems of our societal biases and our emotional irrational desires. This approach opens it up to more critiques; however, it also gives it the opportunity to stand on its own without the two principles appearing as a predetermined goal. In my first chapter, one should get the sense that Rawls’ Original Position is a reasonable way of determining universal principles of justice that govern all of society fairly.

The two principles that Rawls feels govern society fairly are the liberty and the difference principles. Their ordering is important because of a priority of the right, therefore, they are lexically ordered so as to guarantee that rights are prior to the good. (Rawls 1999, 77) The first principle Rawls entitles the Liberty Principle because, it guarantees that every “person is to have an equal right to the most extensive total system of equal basic liberties…” (Rawls 1999, 266) The second principle, the Difference Principle, has two parts that are lexically ordered within it, and balances the equality question within Rawls’ prioritizing of liberty. The Difference Principle arranges social and economic inequalities so that they are to the greatest benefit of the least

3 “The theory of justice may be divided into two main parts: (1) an interpretation of the initial situation and a formulation of the various principles available for choice there, and (2) an argument establishing which of these principles would in fact be adopted.” (Rawls 1999, 47) Rawls chooses to take up his argument over the Original Position in his third chapter, which in my opinion gives the impression that the Original Position is designed to suit the Principles of Justice.
advantaged, and that everyone has an equal opportunity of obtaining offices and positions within society (equality of opportunity). (Rawls 1999, 266)

In the third chapter I will explore the critiques of the larger flaws in Rawls’ theory, focusing on fundamental questions about how the theory works. As the theory is introduced we will begin to see that the principles and the Original Position contain certain inconsistencies and weaknesses. Some footnotes will point to the larger inconsistencies, however I prefer not to pass judgment in the first two chapters rather to provide an ordered analysis of the theory. In the third chapter I will introduce the critiques and present my own opinion on the legitimacy of the critiques. By examining whether the critics make their arguments after using all of Rawls’ theory or not allows me to judge the legitimacy of their claims. Many critiques against Rawls have, for example, not taken into account both principles and their ordering, they have ignored the veil off ignorance in the Original Position, or the have disregarded the ideas of human nature necessary for understanding people in the Original Position. If the third chapter shows where Rawls went wrong then the fourth chapter shows how Rawls fixes his problems.

Rawls attempts to answer his critics in Political Liberalism by presenting a revised look at what he calls “Justice as Fairness.” The two biggest changes that he makes are in how he presents his principles. In Political Liberalism the Original Position is eliminated, and the focus on a moral conception is turned into a solely political conception. The fourth chapter will examine Rawls’ response to the critics before the writing of Political Liberalism with a second response, Justice as Fairness: A Restatement, written in 2001 shortly before Rawls’ death. Rawls also revised A Theory of Justice in 1990, and many of his arguments appear interspersed with responses to specific critiques posed against his theory.
From the critiques and Rawls’ response I will look for two different theories, and beg the question, does Rawls abandon his original conception of justice as a moral theory in his later works? I will draw on Rawls’ responses to criticism and in the conclusion attempt to argue that Rawls looses his argument around the Original Position, but the two principles still maintain certain validity for modern liberal moral theory. The examination will require evaluating how Rawls answers his critics, where his answers are strong and legitimate, and where his answers fail.

In conclusion I will address the future of Justice as Fairness, a conception, and the importance of it in modern society. The conception of justice survives, but the Original Position fails because I believe Rawls initially derived the principles, and then created the Original Position around the principles. The principles are still valid only because, the initial reasons for undertaking his project, an answer to flaws of popular moral theories, are actually grounds enough for the two principles.
The Original Position

My goal throughout this work is to highlight the goals of Rawls in *A Theory of Justice*, and to do this in such a fashion as to advance the Rawlsian argument in a clear concise way. I will be breaking his method down into segments that I find implicit in his way of thinking. The critiques in chapter three should be easier to follow thanks to an “ordered” analysis of Rawls’ theory. Rawls chooses not to present his theory in order of method then principles; rather he switches the ordering presenting the reader with the more pertinent principles first. In order to ease the understanding of Rawls and his critiques I will present his method, The Original Position (OP), first. By presenting the OP first, the two principles deduced from Rawls’ method are much clearer, and hopefully appear as an inevitable deduction by those operating in the OP, something I believe Rawls strives for, but never truly accomplishes.

In order to follow the critiques of *A Theory of Justice*, and the alterations made by Rawls himself in later works I will analyze the OP in stages. The OP itself can be broken down into multiple stages, and I have attempted to separate Rawls’ design into three parts to show the inevitable advancement of his OP towards the two principles of justice. The OP is an all-encompassing hypothetical situation that regulates individual players and leads them to, in Rawls’ view, an inevitable conclusion. The OP has three identifiable segments: 1.) the nature of individuals necessary for one to come to moral conclusions, 2.) the conditions that regulate the negotiations, and 3.) the inevitable method formulated for deriving the principles. Rawls sees all of this as existing in a state of “Reflective Equilibrium” that justifies or champions his principles as being the most appropriate universal ideals.

In order to better see where I am going with my analysis I have broken the three points down into my own hypothetical situation. 1.) First, Rawls believes like Kant that there are
certain characteristics necessary to reach moral decisions. He also holds like Kant that there is a certain cognitive self that is unrelated to external biases. He sets up three traits, self-interestedness, mutual-disinterestedness, and rationality; every human being has these three traits, what I call Rawls’ “Human Nature.”

2.) Secondly, as an individual enters the OP they find themselves restricted to the rules imposed by the position itself. They have all the possible conceptions of justice available to them, but their knowledge is limited to that of general facts about society, they have no specific knowledge about their place or time period in society due to the “veil of ignorance.” Due to the lack of specific knowledge the individual is detached from all their personal societal biases, leaving behind a moral individual with their traits of human nature.

3.) The third and final step of the OP as I have inferred is inevitability,4 because of the rules of the OP, and the human nature necessary for moral decisions that Rawls attributes to individuals under these rules. An individual in the OP is lead through reasoned deduction to choose the maxmin principle (method) for determining the principles of justice. Rawls contends that once the maxmin principle is selected the method will find his two principles as the best configuration of a just system.

Rawls views his two principles of justice as providing the best balance between equality and liberty. The balance will be explained further in the second chapter, however it should be noted that there is an insistence on the priority of the right, and Rawls contends that a theory of justice is based on the goal of obtaining fairness. This involves providing everyone with equal rights and economic benefits, unless a scheme exists where unequal economy benefits even the poorest in society, while at the same time never compromising the provision of rights and liberties to all. Rawls designs his thought experiment of the OP to derive the balance of equality

---

4 Rawls may not agree wholly on this idea of inevitability because inevitability implies that these principles would be come two through some sort of realization or coming to of knowledge rather than through the proposal of principles and ideas which are discussed and either accepted or rejected based on discussion. (Rawls 1999, 129)
and liberty, and to make those involved in determining what is just focus in on prioritizing a system that favors individual liberty, but still preserves some form of equality. The OP finds a just system that lacks nothing but the flaws of the current ideologies.

The Actors

The actors in Rawls’ understanding of what is required for moral decisions to be made leads them to select maxmin as the method for deriving the justice principles. Rawls breaks down the character of individuals as being rational, self interested, and mutually disinterested actors. (Rawls 1999, 129) These three aspects of a player in the OP are the result of Rawls’ reliance on certain Kantian assumptions. Rawls arrives at the three aspects of an individual through examining Kant’s theories and understanding of the self. Rawls provides thorough explanations of why each is of importance to us under the Kantian notion of the self. There are many aspects of Kant that Rawls brings out in his work the three most prominent being autonomy, the categorical imperative, and rationality. (Johnson 1974, 58) The general overarching idea of Kant that Rawls adopts is that there are certain aspects that individuals need to be able to undertake in order to achieve moral principles. Rawls attempts to embody this concept in his OP. There are critiques that get at the three Kantian ideals prevalent in *A Theory of Justice*, but they will be discussed in chapter three.

If we examine the analysis of the OP from the standpoint that we are passing through all the stages it becomes easier to follow and see where Rawls is trying to end up at the conclusion of his work. Rawls begins the OP with a look at the aspects of an individual necessary in deriving moral conclusions he uses the Kantian understanding of the moral self. He believes that there is some moral aspect to individuals that some can possess, and that the aspects are
enhanced or revealed when one is under the conditions of the OP. (Rawls, 1999, 121) This is a Kantian position in the sense that it adopts the notion of a self separate from social interaction, if we examine the dilemma put forth by Kant in a Critick of Pure Reason it becomes evident that Rawls is creating a contract position that utilizes the metaphysics of Kant, but is grounded in a real world hypothetical.

But although all our cognition begins with experience, still on that account, all does not precisely spring up out of experience. For it may easily happen that even our empirical cognition may be a compound of that which we have received through our impressions, and of that which our proper Cognition-faculty...supplies from itself, which addition we cannot distinguish from the former original matter, until long exercise has made us attentive to it, and skilful in the separation thereof. (Kant 1838, B2)

It can be inferred from Rawls that he uses Kant’s statement that there exists this moral individual, a metaphysical understanding of the self. The being that does not “spring up out of experience.” To further this ideal however Rawls uses the veil of ignorance to supplant the “long exercise” necessary for discerning, or uncovering rather, the moral self. Out of all of this metaphysical discussion arises the premise that all individuals in the OP are rational. Rawls sees rationality as a necessity for determining moral principles, “Moral philosophy becomes the study of the conception and outcome of a suitably defined rational decision.” (Rawls 1999, 221)

Included in Rawls’ account of Kant’s metaphysical understanding of how moral principles are derived is the understanding that rationality must be used to come to these moral conclusions. The necessity of rationality in regards to Kant and moral principles is given, but it gains a special aspect in the OP that imparts a new aspect to the basic idea. Rawls uses the “standard definition” of rationality; the actors rank options according to how well certain ones will advance them towards a pre-determined goal. The predetermined goal is one that satisfies
Rawls assumes in this situation that the rational actors in the OP do not suffer from any forms of envy, but rather view inequalities in terms of how just they may be to themselves or to others. Given the definition of a rational individual the other two attributes of an individual can be inferred as inevitabilities, however Rawls does provide us with reasons and definitions for their existence.

The self-interestedness of the one acting in the OP comes out strictly from the definition of a rational actor, and I believe that is inherently implied since a rational actor pursues his own desires he is acting in the interests of himself. Mutual-disinterestedness on the other hand is not as evident from the definition of rationality. It can be inferred that since the actors are rational then they will not be using emotional sentiments to make determinations, and therefore they will not act out of love or passion for others. Rawls has a deeper more involved meaning than the one I have provided; mutual disinterestedness means the moral conceptions of those in the OP should not reflect the motivation of those in the real world. While most people may adopt the general idea of mutual disinterestedness there are benevolent individuals in society who do not do so and for the purposes of constructing the OP it is important to note that not everyone is benevolent. So in a sense the concept of benevolence should be non-existent in the OP, because benevolence might create a situation in which rational players abandon certain desires for the benefit of others. (Rawls 1999, 128) This never actually happens thanks to the conditions imposed on the actors by the OP.
Conditions

There are three main conditions in the OP that form the basic structure, which allows it to exist. 1.) First, all individuals operating in the OP have all the possible conceptions of justice available to them for consideration. (Rawls 1999, 103) 2.) Secondly, all actors in the OP are under a “Veil of Ignorance.” 3.) Lastly, the parties are permitted the ability to enter and exit the OP in order to balance the principles according to the rules of the “Reflective Equilibrium.”

Conceptions of justice are as numerous as the beliefs of people. For the interests of Rawls’ thought experiment he takes on the conceptions that he is critiquing, Utilitarianism, Perfectionism and Intuitionism being amongst them, and attempts to answer the faults in these theories with the two principles that he has derived. Rawls believes that by examining these different conceptions of justice we can eventually derive the features of a just society that are favorable to us as individuals. He states that all we are left to do is determine what scheme will afford us the maximum of these desirable features. It is inferred from Rawls that we will desire the maximum of certain goods because of how he views the individual and the necessary attributes, rationality, self-interestedness, and mutual disinterestedness, for making moral decisions. Maximizing is an important part of the maxmin rule, which is the ultimate method employed by those in the OP to determine the principles. He views our inability to calculate all the possible conceptions of justice as irrelevant in concluding the most just scheme.

The fact that one cannot constructively characterize or enumerate all possible conceptions of justice, or describe the parties so that they are bound to think of them, is no obstacle to this conclusion. (Rawls 1999, 107)

---

5 Even at this early stage it can be noted that Rawls admits to the possibility that other conceptions of Justice exist that might be stronger or more thorough than his own. “Thus although the two principles of justice may be superior to those conceptions known to us, perhaps some hitherto unformulated set of principles is still better.” (Rawls 1999, 106)
What Rawls is conveying here is the fact that there should be a desire to maximize certain attributes in society and given nature of individuals that Rawls views necessary for moral conclusions they will have a goal of maximizing certain concepts to the benefit of all. This means that one must only know the general concepts of justice, and find a system that maximizes those concepts that are viewed as favorable and just. So Rawls can point to certain conceptions of justice that he views won’t fit, and all he must prove is that the conception that he has chosen will maximize rights and liberties necessary in a just society. Therefore, knowledge of all possible conceptions of justice is “no obstacle” to his conclusions.

The Second condition is the veil of ignorance designed to remove social biases from the OP. The veil of ignorance was never originally part of the OP, but upon critiques by Robert Paul Wolff the veil was added to his theory. The veil of ignorance is one of the more important and problematic features of Rawls’ OP. While it removes the individuals from their biases and societal value driven norms, it also has a tendency to create a group of non-distinct individuals. The veil of ignorance nullifies the specific contingencies, which create discord amongst the parties in the OP. It forces the individuals in the OP to evaluate their principles solely on the basis of general considerations. (Rawls 1999, 118) Rawls goes on in §24 to list those aspects of society that specifically relate to a rational individual and are blocked from their knowledge for the purposes of selecting the principles. For my purposes here only the main specific identities need be listed. These particular concepts are all personal social values, the individual’s age, sex, and economic status as well as any other socially imposed forms of identity. Rawls has an important claim about the veil, which is actually an answer to critiques of his OP, but it becomes

---

6 Rational actors given the OP (lacking the veil of ignorance) would find themselves in a standoff between random selection and fair competition. (Choptiany 1973, 148)
very important in analyzing conceptions of justice and choosing those that are most complete. Rawls allows that, “one or more persons can…simulate the deliberations of this hypothetical situation, simply by reasoning in accordance with the appropriate restrictions.” (Rawls 1999, 119) He is pointing out here that the OP is a thought experiment it is not an actual place, rather it is a method for determining whether a society given these conditions is something that rational individuals would find just.

The veil of ignorance permits only general information that affects the principles of justice. In other words, those under the veil will know the basic structure of society, they know all the general laws and theories of justice, and they understand human psychology. Human psychology is important for securing social cooperation by selecting principles that will be readily adopted in society, and will eventually be embodied into a desire to act in accordance with the principles. As noted above Rawls views the veil as a hypothetical situation in which individuals may enter and exit in order to develop the proper universal principles of justice. Rawls calls entering and exiting as necessary for achieving reflective equilibrium. Since, the situation is hypothetical then one can take on the conditions of the OP, develop principles, examine whether the principles are reasonable in the real world, and if not reenter the OP in order to derive more appropriate principles.

Rawls considers the principles of justice to be universal when they exist in a reflective equilibrium. The reflective equilibrium permits us to examine the compatibility of principles selected from all the possible conceptions of justice with the real world outside of the veil. Objections to theories derived through the reflective equilibrium point out that the chosen principles are wrong somewhere. “The important thing is to find out how often and how far it is

---

7 A large complaint arising from this position is that Rawls creates a world in which “objects indistinguishable from one another should be construed as identical for that discourse.” (Hampton, 1980, 334) His rebuttal will raise even larger problems for Rawls’ theory, which I will examine in latter chapters.
wrong.” (Rawls 1999, 45) The OP can then be reentered and a new more appropriate set of principles can be derived. Once we understand how the OP is regulated and who operates within it we begin to see what proposals, conceptions, ideas etc… will work in the OP and why.

For instance the concern over people not being benevolent in the OP is no longer problematic. The rules of the OP create a world where the good of others is taken into account when determining principles. So if one is acting in their own best interests, and they pay no extra due to any other person or group of persons, but they are rational with no knowledge of where they stand in society (veil of ignorance) they will see to it that everyone has their desires fulfilled to the greatest extent. Rawls sums the situation up as:

In brief, the combination of mutual disinterestedness plus the veil of ignorance has the merits of simplicity and clarity while at the same time insuring the effects of what are at first sight morally more attractive assumptions. (Rawls 1999, 129)

Rawls shows that under the OP at this point it become irrational to choose principles that favor any specific group based on their natural advantages. So in Rawls’ view it would appear to those in the OP as irrational to favor a specific skin color, eye color or sex, because an individual in the OP would have no way of knowing if they would in fact be one of the disadvantaged or not once the veil was lifted. There is an aversion to risk created by the lack of specific knowledge it is the last and final attribute necessary before one chooses the inevitable method of maxmin. Rawls views all individuals as having a certain aversion to risk, and this in turn excludes acceptance of arbitrary proposals.

Finally, if the parties are conceived as themselves making proposals, they have no incentive to suggest pointless or arbitrary principles…No one can tell whether such principles would be to his advantage. Furthermore, each such principle is a limitation of one’s liberty of action, and such restrictions are not to be accepted without reason. (Rawls 1999, 129)
Once a player is situated in the OP they are operating as a rational, mutually disinterested, self-interested, risk adverse individual without any knowledge as to their standing in society. From this point one is now required to derive a method for determining the principles of justice, and Rawls argues that the optimal method is the maxmin rule.

Maxmin

The maxmin rule ultimately results in the selection of Rawls’ two principles of justice, because in Rawls’ opinion his principles given the worst-case scenario will result in the best possible outcome. E.G. if one set of principles would result in 2 people receiving $1 and 100 people receiving $1000 that would be unacceptable if a conception existed where 2 people received $3 and 100 people received $500. This is the idea of justice as fairness; fairness is achieved when a conception that benefits the least well off is accepted. When evaluating situations according to the maxmin rule the idea is to rank the worst possible outcomes, in terms of liberty, economic, and social standing; then because of the nature of individuals in the OP they will choose the alternative whose worst outcome is superior to the worst outcome of the others. (Rawls 1999, 133) Rawls does not need those in the OP to know what the actual outcomes are, he just need them to know how they will distribute justice. In simple terms Rawls is attempting to maximize the minimum result.

Rawls outlines three features of situations that give legitimacy to the adoption of this rule. He claims that the OP has these specific features that enable the use of the maxmin rule. 1.) The situation must be one in which the knowledge of all likely outcomes is impossible, thus one cannot compute a monetary gain and use probability to make decisions. The decisions must
be in absolutes, one has to see the principles proposed as acceptable if they were to be adopted, especially in regards to principles that are fundamentally important to others. Probability does not suffice because it makes it easier to take risks. 2.) The individual choosing the principles has little concern for what he might gain above the minimum position that he can be sure of by following the maxmin rule. In a sense the individual will not risk the loss that is possible if they were to strive for a conception that might result in greater gains for a majority because of the possibility of enormous losses to the minority. That individual would not know whether they were in the minority or not, so they would not take that specific risk. 3.) Finally, the nature of the rejected alternatives must be such that their outcomes are unacceptable to everyone, and there are grave risks involved with making the wrong choice. (Rawls 1999, 134)

The person choosing has a conception of the good such that he cares very little, if anything, for what he might gain above the minimum stipend that he can, in fact, be sure of by following the maxmin rule. It is not worthwhile for him to take a chance for the sake of further advantage, especially when it may turn out that he loses much that is important to him… (Rawls 1999, 134)

Rawls is making an assumption that those in the OP will not risk great loss for the advantages provided by larger gains. In the OP under the veil of ignorance it becomes a foolish, irrational idea to chance a slight advantage when the loss could be so great. Approaching it from the standpoint of risk aversion rules out utilitarianism on the basis that one cannot accept a position where even one individual suffers inordinately to the rest of society, because there exists the risk that you might be that one individual.

To examine this principle from the standpoint of the OP it becomes evident that these three features are prevalent. 1.) The veil of ignorance prevents any knowledge of “likelihoods” thus excluding the possibility that one could determine their position through probability calculations. 2.) If it can be established that the principles of justice establish a theory of social
justice that adhere to the demands of efficiency, then the principles represent a satisfactory minimum standard. The purpose of this is to show that all those in the OP can come to mutual agreement, and that outside of the OP the principles will also be mutually agreeable. For instance the prioritizing of the right and the lexical ordering of the principles shows that OP makes individuals less inclined to make economic gains at the expense of liberties, and that it is mutually agreed to under the Maxmin rule. 3.) As long as other conceptions of justice exist that have plausible circumstances that under the OP would be unacceptable, meaning the parties would find the resulting conception intolerable. Given the critiques of utilitarianism and intuitionism in the OP there exist many circumstance under these other conceptions in which unacceptable outcomes could result. (Rawls 1999, 134-135)

In four short pages of §26 Rawls outlines the maxmin rules and clarifies in terms of his OP why the rule would cause an individual to determine that his principles were the most satisfactory. At this point it is appropriate to review what it is exactly that leads one to arrive at the two principles of justice given the human characteristics necessary for determining moral principles and the condition of the OP. I will borrow the method provided by Jean Hampton in his critique “Contracts and Choices: Does Rawls Have a Social Contract Theory?” for the purposes of my review. Hampton begins his critique by breaking the OP into thirteen stages, which he derives from the discussion provide by Rawls in §20. I have taken his technique and changed his analysis to suit my understanding of the OP.

Step one: An individual enters the OP with the goal of determining the most just principles for governing society.

Condition one: An individual making a moral decision, under the Kantian understanding, is rational, self interested, and mutually disinterested.
Step two: The individual thus desires the most primary goods possible.

Step three: The individual realizes that in the OP they are subject to the veil of ignorance therefore they have no concept of where they stand in society.

Condition two: The individual knows that everyone must unanimously agree on whatever conception of justice is chosen.

Step four: The individual also realizes that the circumstances necessary for justice do exist in society. This entails knowing that scarcity is prevalent, and therefore not everyone can have everything.

Condition three: The conception of justice that is chosen will be “final and binding.”

Condition four: All those in the OP are on equal standing with one another.

Condition five: According to Hampton a restriction known, as the “Demands of Reciprocity” must be met. Because one is contracting with individuals that they consider their equals they are forced to adjust their demands and desires in order to reach an agreement. They however will only make adjustments that they know they will be able to live with if accepted into the final conception of justice.

Step five: steps and conditions 1 through 9 lead individuals in the OP to reason through use of the maxmin principle.

Step six: Rawls’ two principles of justice are agreed to because of the three specific features prevalent in the maxmin rule. This means an individual unable to calculate the probability of a conception working (1), and given the desire to gain a maximum from the worst outcome (2), and considering that the risks associated with other conceptions are unacceptable (3) because one would not take unnecessary risks, then the best choice is the two lexically ordered principles of justice. (Hampton 1980, 322-323)
The OP is a thought experiment it is designed to permit people to enter and exit at will. The goal of the OP is to determine principles of justice, however it works in a fashion similar to that of the state of nature used in John Locke’s theory. The OP is a way for one to determine if a government or a governmental action is just. Given the OP one can ask himself would I choose this specific government or would I most likely choose a different conception. The OP applies to the broader moral theories that are used in governmental rule. After entering the OP one follows the steps to their inevitable conclusion, which Rawls holds to be the two principles of justice.
The Two Principles of Justice

Through the use of the maxmin rule John Rawls concludes that two principles would be agreed upon. Those are the liberty and the difference principles. For ease of understanding I have divide my analysis of the principles into five parts: 1.) The general conception of Justice that Rawls believes these principles answer, 2.) the liberty principle involving a detailed definition of what it protects, and 3.) the priority of liberty. 4.) The difference principle, both parts a.) greatest benefit to the least advantaged and b.) offices and positions are open to all under fair and equal opportunity, and 5.) finally the priority of justice over efficiency and welfare. (Rawls 1999, 266) There are really three sections in all: first, the initial view of what is just, understood in the conception of justice, second the principles themselves which achieve that conception, and third the priority rules which lexically order the principles.

The principles themselves serve as an answer to a larger assumption that Rawls believes will be made in the OP, that of a general conception of justice. The general conception of justice that the maxmin rule will bring the actors in the OP to is expressed as:

All social values—liberty and opportunity, income and wealth, and the social bases of self-respect—are to be distributed equally unless an unequal distribution of any, or all, of these values is to everyone’s advantage. (Rawls 1999, 54)

The general conception gets to the heart of Rawls’ concern over society, for him injustice is found in any inequalities that are not to the benefit of all. In the end Rawls is influenced by a desire to develop a moral theory that answers the problems of equality of opportunity.

Rawls’ theory focuses on economic concepts in regards to wealth distribution, and it seeks a situation in which all positions in life are equalized unless an unequal design would benefit everyone. The liberty principle however, does not leave any room for this compromise in
regards to rights. 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. Basic liberties are defined by Rawls as: “political liberty and freedom of speech and assembly; liberty of conscience and freedom of thought; freedom of the person, which includes freedom from psychological oppression and physical assault and dismemberment; the right to hold personal property and freedom from arbitrary arrest and seizure as defined by the concept of the rule of law.” (Rawls 1999, 53)

Liberty takes priority in all moral dilemmas and social settings. Once liberty is prioritized there should always be the largest selection of liberties afforded to all. By prioritizing liberty Rawls creates a realm in which, the only justification for restricting liberties and freedoms arises when it is needed to provide a larger group of liberties, or to relieve worse injustices. (Rawls 1999, 218)

For example I argue that Rawls would accept the suspension of Habeas Corpus during the Civil War as something necessary in a war, which would eventually lead to the restoration of more liberties and rights to everyone. While there may be economic loss on the part of the south when they lose slavery, the economic loss will never trump the restoration of rights to everyone in society. In the end Lincoln was justified in suspending Habeas Corpus because he it was “necessary to forgo part of these freedoms…to transform a less fortunate society into one in which all the basic liberties can be fully enjoyed.” (Rawls 1999, 217) Some may find this example quite contentious, but that is exactly the point there must be great contention involved to justify the suspension of basic rights. The suspension must be legitimated by a desire to see a larger scheme of rights for more people in the near future.
In all other cases Rawls is prioritizing the liberty principle. The priority rules are what lexically order the principles. As I pointed out in the introduction, Rawls is attempting to provide a better alternative to the current moral theories in existence today. Utilitarianism provides a priority of the good, which in Rawls’ opinion can result in the minority being oppressed for the greater good of society.

Thus there is no reason in principle why the greater gains of some should not compensate for the lesser losses of others; or more importantly, why the violation of the liberty of a few might not be made right by the greater good shared by many. (Rawls 1999, 23)

Rawls tries to resolve the danger of legitimized restriction of rights by prioritizing the right in the lexical ordering of his principles. The first priority rule states, “The basic liberties can be restricted only for the sake of liberty.” (Rawls 1999, 266) Rawls views this rule as preventing situations in which the majority is justified in oppressing the minority. He accepts a less than equal liberty only on the condition that those possessing less liberty find it acceptable.

Rawls’ theory in a sense cannot be perverted to cause unwanted pain and suffering through the denial of basic liberties.

Actors in the OP would choose the priority of rights because of the uncertainty of their position, and their inclination to be adverse to risk. The actors have no idea where they will end up in society, therefore they need to guarantee that if they end up in the worst position in society that it will be the best possible outcome that they could have hoped for (Maxmin). Rawls doesn’t truly answer the dilemma of why the worse off individual would prioritize the right, but it appears that it is due to a need for social mobility. Rawls points out near the end of §19 that a sense of Noblesse Oblige will cause those whom are more advantaged to maintain the status quo and not exploit those in lower positions. In turn those less advantaged individuals will see the
inequality and strive for a better position in life. If all positions are open to all, the liberties of those less advantaged are secured, thus providing form more social mobility. (Rawls 1991, 100)

So if one were to be the most disadvantaged in society, they would at least have all the liberties and freedoms necessary to try and advance themselves in society. And in a truly Rawlsian society the second principle would guarantee the ability to advance by preventing exploitation.

Once liberties and rights have been prioritized then it becomes capable for those in the OP to move on and determine what would be the best principle for economic distribution. Rawls state that those in the OP would conclude that the difference principle results in the most just economic distribution. The principle itself has two specific parts that he lexically orders:

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, and
(b) attached to offices and positions open to all under conditions of fair equality of opportunity. (Rawls 1999, 266)

Rawls is attempting to create a society that prioritizes liberty, while at the same time ensures a distribution of income that does not exploit the least advantaged in society. His main goal with this principle is to prevent exploitation of the less advantaged in society. He still allows for inequality, but the inequality must in the end result in a distribution that favors the least advantaged.

In order to apply the difference principle to real social situations or at least think of it as applicable to these situations, we are required to identify the least advantaged group in society. Identifying the one least advantaged individual simplifies the process for determining the least advantaged in society. There are three contingencies that need be identified in order to determine the least advantaged group:
Thus this group includes persons whose family and class origins are more disadvantaged than others, whose natural endowments (as realized) permit them to fair less well, and whose fortune and luck in the course of life turn out to be less happy, all within the normal range (as noted below) and with the relevant measures based on social primary goods. (Rawls 1999, 83)

Rawls points to the aspects, family and class origin, natural endowments, and fortune and luck, as what we must identify in the least advantaged group. He admits though that it is difficult not to arbitrarily identify whether that group is disadvantaged or less well off. He provides one example as choosing a particular social position, an unskilled worker, and then calls this position the least favored. Everyone at or below the income and wealth level of this group is then considered the least advantaged. A more efficient approach to determining the lower strata just involves finding the median and counting all those in the lower half as the least advantaged. “This criterion depends only on the lower half of the distribution and has the merit of focusing attention on the social distance between those who have the least and the average citizen.” (Rawls 1999, 84) Approaching the principle from the standpoint of the median allows for judgments to be made in wealthy countries. Example: the worst off in the United States may appear relatively well off in comparison to those worse off in a third world nation, but approaching the median of each society creates comparisons between wealth and poor in each specific society. Once the least advantaged group is identified moral judgments about society can be made in accordance with the difference principle.

The difference principle holds that once the least advantaged group has been identified then we can identify whether we have a just economic distribution. For Rawls the primary purpose of the difference principle is to provide a just economic distribution in society. He is not saying that the economy should be equal for all, but that given an equal distribution, the only
inequalities permitted are those that raise the lesser advantaged above where they exist in equal distribution. (Rawls 1999, 72) For example if the median level exists at 10 units, but if two people were to be raised up to say 30 units and this results in the worst advantaged (the median level being raised) existing at 12 units then the distribution is just. If the least advantaged were not affected or ended up at 9 units because of the two individuals at 30, then the distribution would be unjust according to the difference principle.

The second portion of the difference principle is the notion that positions and offices are open to all. Rawls has a compassionate understanding of desert that encompasses a notion of social provision for those with less advantageous natural endowments. He views natural capacities as always differing between people, therefore a system that judges people based on their natural attributes is defining them in such a way that does not provide for equal opportunity. “There is no natural feature with respect to which all human beings are equal, that is, which everyone has (or which sufficiently many have) to the same degree.” (Rawls 1999, 444) Rawls is essentially stating that natural endowments or advantages are items that we do not earn or deserve. A naturally intelligent individual no more earned his IQ score than he did the shape of his nose (given that he didn’t have plastic surgery).

So what I believe Rawls would like to do with society given that natural endowments are undeserved is see to it that positions in society are open to all. These positions should not only avoid discrimination, but also be cognizant of the fact that some people have endowments from life that they don’t deserve or didn’t necessarily earn such as their intelligence, their handicap, or their skin color. In the end Rawls wants there to be a focus on equality that encompasses these natural endowments and provides beyond equality of opportunity in order to ensure its maintenance.
A modern idea would be affirmative action; I believe that such a policy given the way in which Rawls interprets desert would be acceptable and even encouraged by the second part of the difference principle. Affirmative action makes up for a disadvantage created by the natural endowment of skin color. Granted in a truly just society the disadvantage associated with skin color should not even exist. Given our society, affirmative action would satisfy the principles of justice by taking into account natural endowments when defining equality of opportunity.

The second priority rule prioritizes justice over efficiency and welfare. It states, “The second principle of justice is lexically prior to the principles of efficiency and to that of maximizing the sum of advantages; and fair opportunity is prior to the difference principle.” (Rawls 1999, 266) Taken at face value we see Rawls is simply ensuring that nothing is taken from the minority for the advantage of the majority. This notion is really the thrust of justice as fairness (the concept).

The priority rule is divisible into two parts; first, the difference principle is prior to efficiency and maximization of advantages, and second fair opportunity is to be provided prior to the difference principle. The first part can be seen as ensuring that exploitation does not become prominent. It may be more efficient or it may maximize advantages to pay disadvantaged persons less and make them work more, but it is unjust according to the difference principle, and would be considered unacceptable by the second priority rule.

The second part ensures fair opportunity, so that if a social scheme were to arise where the least advantaged may benefit economically by an unfair application of opportunity then the scheme would be unjust. If for example the least advantaged in society would benefit if all white men were prevented from working in the fast food industry, this would not be providing fair opportunity, and would be ruled as unjust by the second priority rule.
The last priority completes the over all conception of justice as outlined by Rawls in *A Theory of Justice*. Beginning with the Original Position an individual can use the hypothetical situation to derive a conception of justice. Rawls believes that the conception would involve his two main principles as well as the two priority rules. The conception that Rawls perceives prioritizes liberty and ensures that all economic inequalities are to the benefit of the least well off. It is evident from the design of the principles that Rawls is filling in for gaps created by intuitionism and utilitarianism. He created a set of lexically ordered principles that protect against the exploitation of the minority, and provide fair opportunity to all in society.
Critiques

There are multiple methods for critiquing a work as extensive as *A Theory of Justice*. I could look at it from a strict ideological standpoint entailing general criticism of liberal thought. The theory can also be examined from a methodological standpoint entailing questions into the design of the OP, and the legitimacy of his normative claims about moral decisions. There is an applicability critique, which focuses on the importance of his theory in the real world, and asks can his theory really be applied to modern politics. There is also the critique of the two principles of justice. These critiques tend to focus on one of the two principles and point towards the incompleteness of each principle within the theory. I have chosen to present critiques that focus on Rawls’ methodology, and his two principles because these critiques tend to focus on the theory as I have outlined it here. There are many critiques that focus on modern application of the theory as well as the ideological basis of the theory, however since I have chosen to only briefly address these aspects of his theory in my earlier analysis I have chosen not to focus on those critiques here.

I will begin with the critiques of Rawls’ principles, and how when applied to the real world they may not function as Rawls intends. From there we can address critiques of the methodology employed in deriving these principles. Michael Sandel in his critiques assists in making the transition from principles to methodology; his critique addresses problems of desert associated with Rawls’ second principle, and also questions associated with the legitimacy of the OP. I will begin my analysis with Walter Glannon and Robert Nozick, who attack the difference principle, and question how well distribution will work in a society that prioritizes liberty. Nozick’s line of criticism leads into Sandel and the question of desert, while Sandel’s critiques bring us to the metaphysical problems inherent in the OP. I’ll then conclude with Oliver
Johnson’s critique of the OP and how Rawls incorrectly interprets Kantian ideals into his methodology.

Walter Glannon Critiques A Theory of Justice from the standpoint that there are problems with the difference principle and the second priority rule. Glannon criticizes what he calls Rawls’ worst off priority principle (WOPP), which he derives the principle from the maxmin rule and the difference principle. WOPP revolves mainly around the difference principle and concept that inequalities are only permissible when they are to the benefit of the least advantaged. Glannon reads the maxmin rule into the difference principle and argues that the maxmin requires that the least advantaged be brought up if there is a social scheme available that will bring up these lower groups. The difference is that Rawls believes in an absolute priority of the needs of the worst off, and he does not (unlike utilitarians) take the aggregate effects into account. (Glannon 1995, 427) Rawls lexically orders the advantage to the least well off thus forcing efficiency to take a back seat. Situations arise in where the worst off are brought up through a change in the equilibrium (absolute priority), but a better off group that is not that much better off is injured by the new social scheme. Glannon argues that giving prima facie instead of absolute priority to the needs of the worst off creates a more plausible version of Rawls’ theory. (Glannon 1995, 428)

Rawls bases his theory on providing social goods, and Glannon reads into the purpose of WOPP as attempting to provide the “absolute baseline of sufficiency.” (Glannon 1995, 434) So the worst off must be provided the basic social goods necessary for sustaining life. There should be no comparisons between the wealthy and the worst off. It is not a question of disparity between the wealthy and the poor, but a question of sustainability for the worst off. Glannon presents this definition:
WOPP: In situations of social and economic inequality, we should give priority to the needs of the worst-off social group, regardless of the magnitude of losses to other better-off social groups. (Glannon 1995, 436)

The definition that Glannon provides is one that will take issue with later, however it is important to note that he interprets the maxmin rule as an integral part of the interpretation of the difference principle. He understands that the worst possible outcome that is superior to the worst outcome of others needs to be applied when discussing the difference principle. Since the loss of goods to the better off is of no importance in calculations so the outcome that must be selected according to WOPP is the one in which the least off will benefit. Glannon believes that Rawls’ principle implies an absolute priority for improvement for the least well off.

The strongest case against absolute priority can be made if the better off suffer significant losses and if benefits to the worst off do little to improve their condition…(Glannon 1995, 441)

Glannon provides several hypothetical situations in which the error of absolute priority is clearly articulated. For example take an emergency room full of patients. Provided that WOPP is followed the most severely injured individual will be given the priority of care over all other individuals in the waiting room. It would however make better sense to evaluate those in the waiting room and provide help to those who need care, but at the same time can still be saved. On a battlefield the man who was just shot in the leg is not nearly as bad off as the man who has lost two limbs, but to provide care to the man who is more than likely going to die makes little sense if it is causing further suffering and risking the death of the man who has been shot in the leg. So as a prima facie clause, the rule that we should always provide service to the least well off makes sense, but in the event that the service will cause further suffering to those that are not substantially better off, the concept should be abandoned.

---

8 I diverge with Glannon on the point that it is necessary to apply the maxmin principle when discussing the principles in society.
Glannon derives an altered version of the WOPP that takes into account the damage caused to those in better off positions.

Weighted Beneficence: Benefits to the worst off should be given more weight, but they should not be given absolute priority. For benefits to the worst off can be morally outweighed by a sufficiently large number of benefits to the better off. (Glannon 1995, 448)

Glannon’s weighted beneficence theory takes into account the idea that in order to distribute to the worst off the better off must pay taxes, and therefore the better off incurs a certain amount of loss. If the better off are not substantially better off, and taxation brings those better off individuals below their baseline of efficiency, the worst off create a burden. Since the worst off prevent the better off from maintaining their baseline of efficiency, then the benefit of the worst off cannot take priority.

Glannon also addresses the difficulties of distribution, and the problem created by ignoring efficiency. Since Rawls ignores the number of losses to the better off, he runs the risk of understating the notion of efficiency. Glannon uses the example of a welfare program that supports a very small group of economically worst off individuals. (Glannon 1995, 442) The program place no obligation on those receiving the welfare, and it is supported through a taxation scheme that places a heavy burden on the class of individuals just above the baseline of efficiency. Glannon can see such a situation as leading to a tax revolt, and resulting in less than a marginal increase in the standard of living for the worst off class. Glannon is critiquing such a policy by trying to show “…that a distribution policy may be so inefficient as to lead those who have much to lose, even if they fall above the sufficiency baseline, to question seriously the fairness of the policy.” (Glannon 1995, 442) Glannon proceeds to argue that such a situation would never be conducive to a stable, well ordered society.
Robert Nozick has a similar dispute with Rawls about the errors in distribution. Nozick argues that one cannot maintain a socialist distributive economic system because “liberty upsets patterns.” (Nozick 1974, 160) There are really two things going on in Rawls’ theory: first, the concept of distribution is unjust in the sense that there is no real way of justifying the “confiscation” of another individual’s earned capital; and second, even if distribution worked in a fashion that did not cause civil unrest, the fact that liberty is prioritized affords the ability for equalities to be upset. (Nozick 1974, 163)

Nozick begins his article on “Distributive Justice” as a discussion of how one is entitled to the goods that they possess. He outlines two concepts that in a truly just society repeat themselves with each acquisition. 1.) Original acquisition of holdings, meaning that one must justly come into original possession of their possessions and 2.) transfer of holdings, one must obtain a holding from one individual that justly holds a given item through a voluntary exchange. An example would look like such: If someone comes to acquire something from the state of nature, they must do so in a fashion that does not worsen the situation of others, thus they would fulfill the obligation of a just initial acquisition. (Nozick 1974, 175) The second principle is then fulfilled through voluntary exchanges that are considered just. An unjust acquisition would be one where an individual stole a car and then sold it. While the third person in the chain acquired it in a just fashion, the second person unjustly acquired the vehicle from the original owner, and therefore the third person does not justly own the vehicle. The constant reciprocation of these two concepts creates a just society of exchanges. (Nozick 1974, 151)

Since Nozick uses a scheme of just acquisition in order to justify holdings and desert, it becomes very difficult to justify taxation for the redistribution of wealth. Nozick uses the
examples of a man who chooses to gain pleasure through the ownership of a certain good and the man who gains pleasure through the spare time afforded by the extra nonworking leisure hours.

...whereas the man who chooses not to work the extra time prefers the leisure activities to the extra goods or services he could acquire by working more. Given this, if it would be illegitimate for a tax system to seize some of a man’s leisure (forced labor) for the purpose of serving the needy, how can it be legitimate for a tax system to seize some of a man’s goods for that purpose? (Nozick 1974, 170)

In a system that prioritizes liberty taxation for the purposes of redistribution is the same as forced labor. One man may enjoy going to the movies so he must work longer than the man that gains his enjoyment from watching the sunset. There is no justification in taking the money for the movie from the first man, just as forcing the man that enjoys the sunset to work longer is unjustifiable. I interpret this argument as a critique of Rawls’ difference principle and the fact that the difference principle will inevitably require a redistribution of wealth through heavy taxation. Nozick is pointing to the fact that there is no way of justifying what is taken from those who choose to work and earn more than they need to merely subsist.

The second argument, which addresses Rawls’ notion of nobles oblige, claims that in a society that prioritizes liberty there is no way to prevent people from advancing on their own and earning past their subsistence after a just distribution of wealth has been achieved. Nozick displays the notion that liberty upsets patterns in his example of Wilt Chamberlain. The scenario begins with Wilt Chamberlain a basketball player setting up an agreement that twenty-five cents from each ticket sold at a home game goes to him, and he does it in such a way that each person attending the game has to knowingly place a quarter into a box. Using a box ensures that every person enters into the agreement voluntarily and knowingly. So in the end Wilt Chamberlain makes $20,000 more each year. If the exchange of monies is completely voluntary, then what
justifies taking money from Chamberlain? Better still, how is Chamberlain undeserving of the extra money that voluntary parties are willing to provide him for the pleasure of watching his game? (Nozick 1974, 162) There are two disputes with Rawls that spring from Nozick’s hypothetical. One is what is justified desert under Rawls’ theory, and the other is what in Rawls’ theory prevents the inequalities that will spring inevitably from a society that prioritizes liberty? Michael Sandel answers both of these questions in his critique, Liberalism and Limits on Justice.

Sandel makes specific claims about the nature of desert, and how Rawls conflicts with the position of Robert Nozick. As discussed earlier, Rawls holds that there are certain attributes that individuals do not deserve. Whether it is natural born intelligence, skin color, or social standing, there are certain aspects to life that some people luck out on and others do not. In the end certain people suffer or benefit from natural endowments that they do not rightfully posses. (Sandel 1998, 82)

Sandel uses Nozick’s notion of just exchange to create a theory on how to justify desert under Rawls. If Rawls is correct and the deserts that we obtain through our natural endowments are some how unjustified, then using Nozick’s understanding of justified exchanges makes it difficult to justify the ownership or possession of any item.

If I do not necessarily have to deserve everything I use in producing a thing in order to deserve the thing, what does my desert depend on? (Sandel 1998, 83)

It should be noted that Nozick does not see the problems that Rawls finds in arbitrary assets or natural endowments. Instead Nozick holds that you do not necessarily have to deserve what you have; there are certain things that you may just have and this does not make possession of the natural endowment illegitimate. He does not separate the self and its possessions the way Rawls contends in his theory. (Sandel 1998, 84)
We then have situations where there is almost nothing that any individual deserves. For example, if a child’s parents work hard and use their earnings to raise intelligent hard working children according to Rawls, that child does not deserve to go to the college at the expense of his parents because he has done nothing to earn that desert. If we link Nozick to this concept we can see that the child’s parents don’t even deserve the money that they have to send their child to school because the child never justly possessed their position through fair exchange with their parents.

Sandel argues that we need to have a basis for determining where we stand in society before we can continue with determining desert:

Just as there must be some subject of possession prior to possession, so there must be some basis of desert prior to desert. This is why the question whether someone deserves (to have) his sterling character, for example, is notoriously difficult (for it is unclear who or what is left to judge once his character has been removed.) (Sandel 1998, 84)

In Sandel’s view, once Rawls removes the basic understanding of the self, there are no substantive features of the self-available for us to base desert from. Sandel holds a different view of the person he takes a communitarian viewpoint. The communitarian viewpoint advocates that people identify themselves not as individuals, but as part of a group whether it is by race, sex or social organization. The natural endowment that identify someone as a member of a group then are more than just part of them, but rather one’s identity as a whole. Rawls undermines his argument indirectly by saying not that I do not deserve what is given to me arbitrarily, but rather that I do not posses what is given to me arbitrarily. (Sandel 1998, 84) He is in essence saying that the natural endowments that people are granted by chance do not really make up who they are as an individual, and therefore they cannot posses those endowments in the sense necessary for justifying desert.
On Rawls’ conception, no one can properly be said to possess anything, at least not in the strong, constitutive sense of possession necessary to the notion of desert. (Sandel 1998, 85)

In the end Nozick finds that Rawls leaves us with the Kantian subject that is disembodied from the self. If people are separated from the very things that bind them as humans their identity as a race, gender, or class then they can possess, nothing and become invisible. In the end Rawls is called into question about the autonomous nature that his liberalism claims to embody. (Sandel 1998, 95)

The criticism of desert is part of a larger critique of Rawls, where Sandel looks at the original position and the veil of ignorance. Sandel, focusing mainly on Rawls’ veil of ignorance, constructs an argument around the ultimate loss of self-identity. He takes a communitarian perspective based around how people associate themselves with society and identify themselves as individuals. Sandel sees Rawls as viewing wants and desires, aims and attributes, purposes and ends as what set us apart from each other and what identifies us as humans. Sandel calls this the Humean grounding of the Kantian self.

Each individual is located uniquely in time and place, born into a particular family and society, and the contingencies of these circumstances, together with the interests and values and aspirations to which they give rise, are what set people apart, what makes them particular persons they are. (Sandel 1998, 51)

Through grounding in human attributes, and the metaphysical properties of Kants theory, Rawls derives his theory of justice. Rawls draws a specific distinction between what one possesses and who one is. You are related to your traits and desires in the fashion that they are yours not mine, but you are distanced from them in the sense that they are yours, not you. (Sandel 1998, 55) The distancing is what Rawls believes signifies the Kantian understanding of

---

9 The very subject that Rawls sought to avoid in his theory.
what is necessary for moral decisions. If one can separate what is mine and still understand who
they are then there is something more fundamental than traits and desires. There results what
Sandel calls a duel aspect of possession in which what is mine eventually becomes me.

As the desire or ambition becomes increasingly constitutive of my
identity, it becomes more and more me, and less and less mine. Or
as we might say in some cases, the less I possess it, and the more I
am possessed by it...As it grows from a desire into an obsession, I
possess it less and it possesses me more, until finally it becomes
indistinguishable from my identity. (Sandel 1998, 56)

The idea that we are what our desires and traits say we are makes the Original Position a
difficult area to derive contractual agreements. If I identify more with my skin color or gender,
and in the OP under the veil of ignorance have no knowledge of this identity, how can I make
moral judgments? I cannot even obtain a full understanding of my self? The communitarian
understanding of society is that self, the possessions that constitute who I am. The natural
endowments that make me me appear to be what people identify themselves with, rather than this
separately constituted individual. They are also necessary for deriving moral principles. For
Sandel, if man views himself in this fashion, then the veil of ignorance will prevent discussion or
contractual agreement in the OP.

In the book Groups and Group Rights, a collection of essays discussing the concept of the
group in liberal democratic thought, many of the scholars discuss the nature of the group and
how an individual relates to the group. They continually reiterate Sandel’s communitarian
conception in which the individual ceases to view themselves as individuals, but rather they
identify with their natural endowments that group them with others in society. Some of the
strongest group bonds involve what Rawls constitutes as the natural endowments, those
undeserved traits of man.
If we focus on culture here, it is because it is a pervasive source of social identity, providing a context for thought and action that involves language, values, modes of behavior, education, socialization, practices, traditions, and shared history. (Sistare 2000, 46)

The types of group membership that Sandel and other scholars are discussing are involuntary groupings. Unfortunately for many groups, liberal theory focuses on the individual and group rights are left out in the resulting contractual agreements. (Sistare 2000, 137-139) In a sense when individuals are stripped under the veil of ignorance they have no understanding of how they relate to a group. Rawls makes a fundamental assumption that there is a unique individual within everyone that is not constituted by outside influences. Group rights argues that people are constituted by these outside influences, and that there is no individual self separate from all these traits. Rawls especially misses any opportunity to include the group identity within his overall understanding of the OP because of the veil of ignorance.

Sandel concludes his critique of Rawls by pointing to the lack of a contractual agreement behind the veil of ignorance. There can never be a fully constituted individual within the OP and therefore there is no need for more than one person. Rawls contends that there is a contract, but a contract requires discussion and agreement. In the OP, there is only a “single subject” with the attributes necessary for making a moral decision. The justification for the OP lies not in what individuals agree to, but rather what they apprehend. Individuals in the OP discover the principles, and have no real room for discussion about what is concluded. For Sandel the OP “is not a contract after all, but the coming to self-awareness of an intersubjective being.” (Sandel 1998, 132)

The end result of Sandel’s argument is that Rawls has no way of arriving at a contractual theory, and there is a weak foundation at best for the principles that are derived. Jean Hampton
follows the same line of argument in her critique of the original position. She contests the idea
that the OP is a contract, and provides an alternative method, advocating that Rawls enjoys a
very Kantian metaphysical methodology. A method in which there may be a better constituted
individual in the end contract, but in the end his critique finds that there is still no contract
because of the ultimate reliance on the maxmin rule.

In his article, *Contracts and Choices: Does Rawls Have a Social Contract Theory?*,
Hampton discusses how in the OP there is no contract resulting in an agreement on the two
principles, rather there is a separate procedure that results in a contractual agreement. Hampton
admits that it is necessary for those in the OP be objective and impartial, however she believes
the veil of ignorance obscures too much of the individual, the same complaint of Sandel.

…we get a very peculiar selection environment where there are no
conflicting interests that need to be mediated, where everyone
prefers the same two principles, and where the agreement on there
principles is unanimous. (Hampton 1980, 316)

The veil of ignorance removes all of those attributes that separate people into individuals
and in a sense creates being with no sense of self. The dilemma then becomes how do people
reach agreement or achieve a contract when they all have the same cognitive processes.
Hampton outlines a second procedure, which she finds in part one of *A Theory of Justice*. She
believes that this second procedure will allow for a contract or agreement to be derived.
(Hampton 1980, 315)

Hampton sees two distinct methods employed by Rawls. First, Rawls uses the idea that
there is a constraint placed on all actors in the original position that the principles are final and
binding (the Irrevocability Constraint). Hampton sees the constraint in conjuncture with the veil
of ignorance as forcing the acceptance of the principles on the basis that those in the OP would
get a decent share of social goods given the possibility that they would find themselves in the
lowest social strata. The second method employed does not require the veil of ignorance, instead the irrevocability constraint combines with what Hampton calls the "restriction from the demands of reciprocity." The demands of reciprocity creates a constraint that requires those in the OP to take into consideration the interests of others as well as one’s own interests. (Hampton 1980, 322)

The first method employs a controversial claim by Rawls about the nature of individuals. Rawls makes the claim that once in the OP the normal aversion to risk is increased.

(H)e maintains that they have a normal risk aversion, but that, because of the extensiveness of the losses they would suffer if they were the least advantaged in a society, they have good reason to be conservative. (Hampton 1980. 321)

So merely because they could have the possible misfortune of being the worst off in society, and the veil of ignorance makes it unknown whether they will be worst off or not, the actors choose the best worst case scenario. Normal risk aversion rationalizes in the OP a stronger aversion to risk than what exists outside of the position. The result is not a contract, but rather a result of the veil of ignorance and the fact that there is a chance that one loses much that is of value to himself if he risks his options.

The second method interpreted by Hampton involves the restriction from the demands of reciprocity, which directly affects the way individuals in the original position come to a conception of justice. The method employed results in the exact same principles, however, it results in a contract rather than a coming to self-awareness. The new method relies much less on a need of the veil of ignorance, and more so on the reasoning of multiple individuals.

Each party in M2 (method 2) is concerned not so much with his possible identity in society as with the pressure other parties in the original position will exert upon him to accept a conception of justice that will distribute primary goods fairly among them. (Hampton 1998, 323-4)
Hampton is making a much larger argument than simply that Rawls provides for two methods of deriving the principles of justice, and that each utilize different parts of the OP. Hampton’s larger argument is that Rawls need not contractually justify his principles, and in fact the principles of justice cannot be contractually justified. In order to determine whether method one or two truly allow for a contract Hampton derives a two-part definition of contracts. First, agreements are distinguished by parties that grant promises to one another, therefore each of the actions undertaken by the parties is considered binding unless the parties both agree to release each other from the commitment. Secondly, a contract involves some type of exchange between parties; it is a “two-sided affair.” (Hampton 1998, 324-5)

The failure is that in the end the principles are derived using the maxmin rule. No matter how the OP is analyzed, no matter which method is used to derive the principles, the principles are ultimately derived through a realization that the maxmin rule is the best way of deriving moral principles. Hampton uses game theory to point to the maxmin rule as a method used not in group decision making, but rather by individuals in situations where there is no certainty of outcome. The maxmin rule is used when the decision maker is completely ignorant of their position in life, and when they have no idea of what their particular circumstances will be in the future. (Hampton 1998, 326)

The veil of ignorance places the parties in the exact position where they are forced to select the maxmin rule as the best method for deriving the principles of justice. The veil unfortunately creates problems for those individuals in the OP, as mentioned before, by making it impossible advocate different conceptions of justice.

Instead, all of them, after reflection prefer the same conception, because all of them are faced with the same problem of uncertainty, leading them separately and independently of one
another to invoke the maxmin rule in order to solve it...There is only a series of independent but identical deliberations and decisions in this “uncertain” selection environment. (Hampton 1998, 326)

The veil of ignorance in the OP creates an uncertain situation where even method two is forced to use the maxmin rule. However, if the veil were lifted, it can be seen that parties in the OP would deliberate in such a way as to derive the two principles through use of the demands of reciprocity. Once the veil is lifted all of the parties are then subject to the same prejudices that make agreement all but impossible, and it becomes very unlikely that Rawls’ two principles could be derived. Hampton argues that Rawls in fact admits in a Theory of Justice that deriving the two principles from an actual contractual agreement is all but impossible. The end result is that the veil of ignorance is necessary for the groups to reason from the maxmin rule, and at the same time the veil prevents the possibility of an actual contract being derived. (Hampton 1998, 328)

Hampton’s final point of contention is with the lack of finality in the OP, Rawls never actually states that there is finality to the decisions made. There is a final portion of the OP that is the irrevocability constraint, the idea that the process itself cannot be repeated. Hampton lays all of the arguments out on the understanding that principles are final and binding. However, it is not true that the principles are final because Rawls himself calls the OP a hypothetical situation in which new conceptions of justice can be tested. Hampton concludes that Rawls is not arguing for a social contract, but rather for a “Kantian device” to legitimate his political theory. (Hampton 1998, 336)

As articulated earlier Rawls organizes his OP as an interpretation of Kant’s autonomy and categorical imperative within an empirical theory. If one sets out to pursue the moral law, then that individual will pursue only those conceptions that this form of reasoning dictates. By
acting prior to the definition of the moral object, the agents guarantee that their desires do not become problematic in the selection process. For Kant placing reason before the selection of moral objects allows the process and the individual to be completely free. Hampton sees what Rawls is doing as purely Kantian and not really a social contract. Rawls’ OP actually has autonomous individuals testing which conception is right by considering what it would be like if everyone acted in the same fashion, and he also uses a method of reasoning to derive his moral principles. Both of these points are in keeping with Kant’s method for determining moral theory. (Hampton 1998, 336-7)

In conclusion Hampton argues that Rawls’ OP is better understood from the standpoint of one individual using practical reason similar to that of Kant’s, rather than a group of different parties debating and discussing conceptions in order to forge a contract. If we accept Hampton’s critique that Rawls’ OP leaves no room for a contractual agreement, then in order to maintain the OP we must accept Johnson’s alternative conclusion that the OP is truly a Kantian conception.

If we can show that on top of not being contractual that the OP is not Kantian, then there is some real question as to the neutral grounding of the OP. Oliver Johnson provides us with a critique of the Kantian nature of the OP. He claims that Rawls misinterprets Kant in §40 and misrepresents the categorical imperative and the notion of autonomy.

In “The Kantian Interpretation”, Johnson argues that Rawls must radically misinterpret Kant in order to fit his views in with those of Kant. Johnson approaches Rawls from three closely related points found in Kant, mainly (1) autonomy, (2) the categorical imperative, and (3) rationality. Rawls claims to employ all of these points in section 40 of A Theory of Justice where he discusses the Kantian interpretation of his theory. (Johnson 1974, 60)
1) Autonomy according to Kant is when our motives are born out of respect for the moral law. With autonomy the emotional desires of man play no role. Johnson describes autonomous actions as those associated with a sense of duty. He points out that in Rawls the players do not decide on a conception of justice autonomously because of particular “considerations or contingencies,” but rather because of the veil of ignorance. (Johnson 1974, 61) The individuals in the OP do not choose the principles out of a sense of duty. Instead, they act autonomously because of the veil of ignorance imposed upon them. Johnson argues that the veil of ignorance cannot simply be placed upon the individuals and make otherwise heteronomous decisions autonomous. The individuals must actually choose to reason from moral law and view it as their duty.

2) The categorical imperative states that a decision is made without a given purpose in mind. Kant distinguishes the categorical imperative from a hypothetical imperative, one in which the individual is motivated by a desire to act towards a specific end. Johnson argues that Rawls misses the distinction. Rawls states that a hypothetical imperative is one where an individual acts according to specific ends, and when one acts with a desire for a “primary good,” goods all humans seek, that this is then a categorical imperative. Instead Johnson states that Rawls is distinguishing between “rules of skill” and “counsels of prudence,” which are both forms of a hypothetical imperative. (Johnson 1974, 63) Even if it is a primary good that motivates the actors in the OP, the primary good is still a specific end.

3) Rationality is defined by Rawls as making choices that will advance the best possible ends of the individual making the decision. Johnson finds an error in Rawls’ conception because if the actors in the OP are only rational to the extent that their decisions serve their own long term best interests then that reason is strictly instrumental.
Reason…is not functioning as a moral faculty…if it is being used as an instrument for the satisfaction of our various desires…On the contrary, the moral and proper function of reason (according to Kant) is to produce a will good in itself. (Johnson 1974, 65-6)

The definition of rationality raises serious faults as to the morality of Rawls’ theory. If Rawls is using rational choice instrumentally, then according to Kant Rawls’ theory is not moral. If Johnson is properly applying Kant to Rawls then §40 in which Rawls discusses the Kantian nature of the OP is a misinterpretation of Kant’s theories. Therefore Rawls’ OP is not really a moral situation according to Kant. If we accept the argument that the OP is not a social contract and Johnson is correct that Rawls misinterprets Kant, then the OP has no grounding. If the OP has no neutral grounding the principles of justice cannot be situated according to a specific methodology, and therefore fall into question.

The question raised by all of these critiques is if the OP is not contractual and not Kantian. If the OP is invalid then the moral grounding of the principles falls into question. It begins to appear that the OP was created specifically to derive the original principles, and therefore does not serve as any sort of grounding at all.

The critiques in the chapter above cover a Theory of Justice from the two principles to the initial Kantian conception of the OP. There was so much discussion around his theory that he was eventually forced to write a response, Justice as Fairness: a Restatement, to the critiques that found serious error within his theory.
Rawls’ Response

There exist so many critiques of *A Theory of Justice* that Rawls was forced to respond and clarify his original thesis. Eventually, he changed his thesis in its entirety, and published a work entitled *Political Liberalism* a piece that continued to grant priority of the right however it did so through what Rawls called an overlapping consensus. In his work *Justice as Fairness* Rawls advocates an overlapping consensus, which makes the work less useful when attempting to defend the OP. However he still provides the readers with a defense of the original principles. He never changes his position on the principles of justice themselves, just the method for arriving at the principles.

I have not defended the accuracy or correctness of Rawls’ theory throughout this work and will not attempt to defend him at this stage either. I will be interpreting how Rawls would respond to the critiques based off of my own interpretation of his writings. I believe that the OP was not developed strongly enough to fend off the critiques I have presented thus far. The basis for the principles is actually stronger than the OP; a second basis for the principles can even be derived from Rawls’ Utilitarian critiques. If the OP is not valid then it represents a construction for the sole purposes of deriving the principles that Rawls advocates. The Principles are not invalidated however just because the OP is not neutrally grounded.

Rawls does have certain points within his theory that are defensible, but much of his method is difficult to defend even for Rawls. I hope to show that the critiques focusing on the OP cut deeply into the reasoning of Rawls’ theory, however the critiques of the principles do not carry as much weight. By using the revisions in *A Theory of Justice* and *Justice as Fairness*, I will derive direct responses to the critiques outlined in chapter three. The quality of each response will depend upon the ease with which his theory can be defended.
Walter Glannon’s critique of the difference principle is utilitarian; he advocates that there is a certain loss of efficiency with Rawls WOPP, because there exists a possibility that the greatest good might not go to the greatest number. If we examine his final alternative of weighted beneficence, the outcome is incredibly utilitarian in the sense that the worse off can only benefit if their benefit does not cause more people to suffer. In the end what Glannon qualifies as a more efficient alternative theory is really a more utilitarian solution. His alternative ensures the greater good is provided to the largest number of people.

Utilitarianism was the very moral theory that Rawls found flaws with and sought a solution to. His theory was initiated with the goal of prioritizing the right. Therefore Glannon’s argument is not really a critique but a reiteration of the same faulty utilitarian principles that Rawls is attempting to correct. His critique fails because Rawls is purposefully not supporting the utilitarian ideals.

Glannon’s alternative represents the exact same faults of utilitarian theory that Rawls was trying to resolve. Using one of Glannon’s analogies, there becomes justification for leaving the man with more severe injuries in the ER waiting room, or keeping the truly poor in hovels because it would not, in the end, benefit the most people to provide service. Glannon does not provide a fully accountable moral theory, because it still allows for minority groups to be discriminated against based on the benefit of the majority. Rawls easily disputes a utilitarian critique based on the point that it can never be acceptable to discriminate, however harshly, against a small minority group for the benefit of the better off larger majority.

In *Justice as Fairness* Rawls outlines an efficiency graph, which is reproduced below. In this graph he outlines a production curve (Output of Production, op curve), and a justice
conception curve. He then points to where utilitarianism falls on the graph in comparison the difference principle.

\[ y = LAG \]

\[ x = MAG \]

(Rawls 2001, 62)

The Graph above outlines three different conceptions of justice, and points to where they fall in relation to the most and least advantaged within society. The y-axis represents the income of the least advantaged while the x-axis represents the salary of the most advantaged. Everything below the forty-five degree angle, which represents equality, denotes equal justice lines. “They are equal justice lines in the sense that any point on a line is equally acceptable provided it is reached by an op curve of a scheme of cooperation satisfying the principles of justice prior to the difference principle.” (Rawls 2001, 62) Point N represents Nash where the product of utility is maximized, point B represents Bentham where the sum of individual utilities is maximized, and point D is the point closest to equality that still maintains efficiency. Point F is where a feudal society would fall along the op curve.

Any point along the op curve refers to the income of the least and most advantaged. any point along the curve then corresponds with a conception of justice. “One scheme is always more effective than another if its op curve always gives a greater return to the less advantaged for any given return to the more advantaged.” (Rawls 2001, 63) The Highest point on the op
curve represents the most effective scheme, therefore, the difference principle qualifies as the most just and efficient distribution of justice because it directs society toward that high point.

The Bentham point represents the alternative scheme proposed by Glannon. We can see that Glannon’s scheme is actually closer to a feudal system rather than an efficient just design. According to the Bentham point the most advantaged group would benefit greatly compared with the least advantaged group. Glannon’s argument is actually not efficient according to the op curve, and fails by not attempting to equalize a just scheme close to the high point on the op curve. In the end the greater returns of the more advantaged will be used in Rawls’ scheme to cover the costs of training employees, providing incentives, and to mark positions of responsibility. The difference principle is represented by point D on the graph, and marks the largest return to the least advantaged. Therefore the Least advantage benefit as best they can without bringing any of the better off below the base level of subsistence. (Rawls 2001, 63-64)

While the graph shows how Glannon’s critique is a utilitarian conception that will benefit the most advantaged. I believe that since Glannon’s critique is utilitarian it suffers the most because it allows for a less advantaged minority to be oppressed on the basis of good for the majority. While the difference principle may disadvantage some more advantaged individuals, it becomes very difficult to support a scheme that forces the least advantaged to maintain their position. Moral theory implies that one is doing the “right” thing; one is trying to visualize a system that benefits everyone to the greatest extent. It is not a sufficient moral theory when the suffering of a minority can be justified based on the need to provide the most people with a standard of living above the baseline.

Nozick disagrees with the utilitarian argument as strongly as Rawls however he takes a very libertarian position asking how one can justify the taking of another’s earned position in
life. Sandel, who is much less of a libertarian than Nozick, critiques the justification that Rawls provides for a tax scheme that would redistribute wealth. Nozick interprets Rawls as justifying taxation on the basis that many aspects in life many deserts are undeserved, the notion of desert working in both a positive and negative fashion.

I believe Rawls would argue an alternative conception of desert; one separate from the interpretation of Sandel and Nozick. He would propose that we leave moral desert to private individuals, and what we really need to look for is a replacement that works as a “reasonable political conception.” Rawls proposes the idea of legitimate expectation. He believes that when one acknowledges undeserved differences, people will become more competitive, and will begin to educate their endowments. By acting on natural endowments and working to improve them, the action of working on the endowment, studying, exercising, working, etc… is something that an individual does have a claim to. These endowments then become part of the individual and are deserved by the individual. (Rawls 2001, 73-74) Rawls does not really repute Sandel’s critique, rather he provides an alternative understanding of desert; one in which natural endowments may not be deserved, but can be earned.

Sandel has misinterpreted Rawls in his critique of desert. Rawls does justify desert, but he does so through a clause that endowments need to be earned before a desert is deserved. One can earn desert through working on natural endowments and improving them. There is no way to guarantee that one succeeds in society thus we allow for competition to exist. Everyone can enjoy their endowments and make a certain claim to desert as long as those in society with advantageous endowments use them to benefit the greater good of society. Those with fewer advantages need to work to improve their position, and society will provide for them and grant them certain leeway in an attempt to legitimate competition.
I find it much more difficult to reply to the critiques around the OP in such a way as to continue its basis as a strong moral methodology. Sandel points out a serious error associated with the OP, and Oliver Johnson’s critique of the Kantian aspect of the OP is also very difficult to rebut. I really see Rawls himself abandoning his method of the OP and opting for the idea of an overlapping consensus. He maintains his principles, and his initial notion of justice as fairness. When it comes to the OP however there is a need for more solid understanding and Rawls it appears fails to provide a basis whether it is in Kant or in a contractual conception.

The brunt of Sandel’s critique is communitarian in nature, and focuses on the notion that the veil of ignorance creates a realm where one needs to accept the notion that all individuals view themselves as individuals, and not part of some larger group. Rawls however does not contend that all individuals need to view themselves as rational, self interested, mutually disinterested beings. He claims that these are the aspects necessary for making moral decisions, whether one identifies with a group or, as an individual has no bearing on the moral conclusions that provide for society. However I will acknowledge that it appears to still be a constraint, even if it is a constraint over the aspects of a moral decision. A good rebuttal to Sandel is to really look back at the principles themselves. The liberty and the difference principle, if they are good moral principles, should provide for groups as well as individuals. To claim that one could not derive the principles given the nature of the OP is a different critique. Jean Hampton’s critique of the OP as a social contract really gets at a question over Rawls’ methodology.

Rawls would most likely agree with Hampton that his theory is not a contract. He would say that there are certain groundings in contract theory, but that in order to fully understand the OP one needs to look a the Kantian conceptions.

The original position may be viewed, then, as a procedural interpretation of Kant’s conception of autonomy and the
categorical imperative within the framework of an empirical theory. The principles regulative of the kingdom of ends are those that would be chosen in this position, and the description of this situation enables us to explain the sense in which acting from these principles expresses our nature as free and equal rational persons. (Rawls 1991, 226)

The Kantian understanding of the OP is more prevalent that the contractual design, and Hampton is correct in claiming that the OP does not really derive a contract. Rawls is very ambivalent as to whether there need be a group or whether a single individual can enter the OP and derive the principles of justice.

Some may object that the exclusion of nearly all particular information makes it difficult to grasp what is meant by the original position. Thus it may be helpful to observe that one or more persons can at any time enter this position, or perhaps better, simulate the deliberations of this hypothetical situation, simply by reasoning in accordance with the appropriate restrictions. (Rawls 1991, 119)

Because of the ambivalence towards the number of individuals in the OP it is far too difficult to justify a contractual conception. The Kantian methodology is easier to apply for individuals given that the situation is “hypothetical.” Oliver Johnson’s critique builds off of the Kantian methodology and argues that Rawls completely misinterprets the intentions of Kant. Stephen L. Darwall defends Rawls, and provides an answer to the contentions of Johnson.

The mistake that Johnson makes is that he interprets autonomy in the OP with respect to the principles. Darwall contends that the principles are chosen by those in the OP so as to promote the interests of each party. Since the decisions in the OP are motivated by the interests of each party they are heteronomous. Johnson thinks that in order for the principles themselves to be autonomous the choice in the OP must be autonomous. Darwall states that the autonomy that Rawls advocates comes in the decision to enter the OP. (Darwall 1976, 166) To act under the principles in the OP is an autonomous decision because of the veil of ignorance.
Thus, if a rational being chooses to act on principles which would be acceptable to him if he were under the veil…such a choice is by no means a choice on the basis of his interests and thus is not, on those grounds, a heteronomous choice. (Darwall 1976, 166)

Rawls is not claiming that the principles are derived autonomously, but rather that when one acts on the basis of principles acceptable from the hypothetical OP then one is acting autonomously. It is really a matter of when one acts on the principles, whether one is in the OP or outside of it that determines whether the actions are autonomous. Rawls sees the OP as a procedure through which one acts autonomously in selecting the principles of justice. Therefore the OP is entered into autonomously and following the derived principles is autonomous, while the choices under the veil might be heteronomous. (Darwall 1976, 167) Darwall argues that Johnson has the same argument for all three Kantian ideals, and only briefly rebuts the notion of rationality and the categorical imperative. “Johnson makes the mistake of supposing that since something characterizes the parties’ choice of principles within the original position…it must therefore characterize the principles…” on the grounds that an actual person grounds the principles in the same fashion. (Glannon 1976, 170)

Rawls is easily defended on the basis of his principles. The methodology however is much more difficult to defend. When we get down to the Kantian understanding of the OP it is much easier to support the OP however the contractual aspects have too many pitfalls to justify the existence of the OP on its own. The principles themselves exist as a stronger argument only because of what they convey as an ideal. The reasons that Rawls laid out for determining a scheme of justice that prioritizes the right are in and of themselves moral. He sees the suffering of a minority group unacceptable, and argues for a conception of justice that avoids that suffering if at all possible. The OP used to arrive at such a conception is much more difficult to support only because of the abstract nature of the position. Rawls himself is ambivalent towards whether
the OP requires more than one person or not, and whether it is in fact a contract or contract like situation. I believe the reality of it is that the OP needs to be interpreted as a contract like conception that attempts to ground moral conceptions according to the metaphysical concepts of Kant. If his OP is interpreted in this fashion many of the critiques that say it isn’t truly a contract or it isn’t truly Kantian are nullified, because the OP is acknowledged as being neither. If the validity of the OP can be maintained then the legitimacy of the two principles can also be maintained.
Conclusion

I do not really see the OP holding up against the arguments over it being a social contract and the misinterpretation of the Kantian understandings. I have to agree with Hampton’s critique that the OP cannot be a social contract. The Veil of ignorance appears to take away too much of a person’s identity, especially the parts that would be necessary for arriving at a moral conception. If someone has no knowledge of where they stand in society then there is no way they can arrive at a conclusion that draws on all of their social interactions.

The Kantian critique while slightly more complex follows a simple line of thought. I have to agree with Johnson’s argument that no matter whether you exist in the OP or outside of the OP the concepts proposed by Kant need to be followed. Darwall’s rebuttal appears to hinge on a technicality of whether one is in the OP or not. To say that the choice to abide by the conceptions coming from the OP is autonomous does not really answer the problem of the OP and the principles coming from it not being autonomously chosen. The agreement to abide by the principles being autonomous is not the same as the principles and the OP being autonomously chosen by the individuals. If the OP does not really follow the autonomy in Kant or rationality and the categorical imperative for that matter, then the OP cannot really be considered Kantian.

To argue that the OP is a methodology grounded in contract theory, and utilizing Kantian ideals is invalid. I cannot agree with the OP if it is so easily disproved as a. a contract theory, and b. a Kantian based methodology. I do not believe that the Principles themselves are made invalid by this realization, and my ultimate conclusion is that we can still have the principles of justice even though the OP has been knocked down.
I believe that the inadequacy of the OP is a result of the fact that the principles came first and the OP was really derived as a method for legitimating the two principles. The principles were not created out of the OP in reality I believe Rawls derived the principles as an answer to utilitarianism first. If we look just at the faults found in utilitarianism we see that the principles really do answer and solve the faults found in the moral theory. The two principles guarantee basic liberties to everyone and therefore protect against the oppression of small minority groups. The difference principle advocates what I see as another important part of moral theory, and that is the elimination of exploitation, yet what I see as the continued maintenance of a capitalist free market through the available inequalities.

My final position is that while the OP is not strong enough to stand up against criticism, the principles still can exist as legitimate moral conceptions. I can visualize a government that abides by the original principles, and exists as a stable well balanced secure society. A Theory of Justice may have its flaws, but the basic premises are derived from legitimate concerns over a lack of certain guarantees in Liberal theory. The two original principles really answer the faults of many modern moral theories, and can be interpreted as a base for a strong society.
Works Cited


