
JUSTICE AND THE PUBLIC SPHERE: A CRITIQUE OF JOHN RAWLS' POLITICAL LIBERALISM

Wanpat Youngmevittaya

University of York, UK

ABSTRACT

This article criticizes John Rawls' conception of political liberalism, which insists that political sphere governed by his two principles of justice can be separated from any comprehensive moral doctrines, and that the validity of his conception of justice is political, not metaphysical nor comprehensive. I argue that Rawls' project is flawed by showing that his two principles of justice and political liberalism are presupposed by the very comprehensive/metaphysical doctrines which he denies. Whether he realizes it or not Rawls chooses a particular comprehensive theory of the good/person, specifically that of an unencumbered self. I discuss Rawls' political liberalism from two points of view. First, I discuss Rawls' political liberalism from political economy points of view, which I argue that the foundation of Rawls' principles of justice lies in his particular theory of the person. Second, I discuss Rawls' political liberalism from philosophical points of view, which I argues that Rawls' political liberalism and theory of the person are comprehensive, and that political sphere cannot be separated from private sphere.

Introduction

This article aims to criticize John Rawls' political liberalism, which proposes that his principles of justice governing public sphere are derived from none of any comprehensive moral doctrines so that they should be acceptable to 'everyone' whose moral comprehensive doctrines are different; it is not only possible, but also ought to separate justice from any comprehensive doctrines. In other words, he claims that his principles of justice are justified on the basis of a political consensus, which is neutral among competing conceptions of comprehensive doctrine. I will argue that Rawls' political liberalism is flawed in the sense that his principles of justice are far from being neutral, and that it is *impossible* to separate political sphere from any comprehensive doctrines as he claims. The political sphere requires a judgment of comprehensive moral doctrines instead of the claim of neutrality between them.

This article consists of three main parts. First, I give an outline of Rawls' political liberalism, which I link his *Theory of Justice* (1971) with his *Political Liberalism* (1993), and other works such as *The Law of Peoples* (1999). I argue that Rawls' project is to preserve Kantian morality without referring to any comprehensive doctrines as Kant did. Second, I discuss how Rawls' principles of justice are really justified from the perspective of political economy, especially from works of Harsanyi (1975; 1977) and Hampton (1980). In this section I show that the secret of Rawls' principle of justice is not based on the social contract, but on his particular theory of the person or the circumstances of justice in the original position. Finally, I discuss how Rawls' theory of the person is comprehensive rather than political, in which I argue that it is comprehensive because it is justified on the basis of rationality which one needs to assume what characteristics can be called rational and good or just in the first place.

Rawls' Political Liberalism: An Outline

In *A Theory of Justice* (1971), Rawls' powerful claim is that any theory of justice governing the whole society *must* be acceptable to 'everyone' in the society. One of the most important tasks of Rawls (1971) is to defeat utilitarianism, the dominant theory in economics and political philosophy at that time, in defining "justice" for the society, which argues that any just principle governing the whole society must maximize the greatest happiness of the greatest number (Bentham, 1879; Harsanyi, 1975; Sandel, 1994; Swift, 2014). According Rawls, utilitarianism is not a proper principle of justice because its violation of the principle of the separateness of persons; it prioritizes the welfare of society as a whole over individual rights, as Rawls (1971, pp.3-4) argues:

"Each person possesses an inviolability founded on justice that even the welfare of society as a whole cannot override. For this reason justice denies that the loss of freedom for some is made right by a greater good shared by others. It does not allow that the sacrifices imposed on a few are outweighed by the larger sum of advantages enjoyed by many. Therefore in a just society the liberties of equal citizenship are taken as settled; the rights secured by justice are not subject to political bargaining or to the calculus of social interests. The only thing that permits us to acquiesce in an erroneous theory is the lack of a better one; analogously, an injustice is tolerable only when it is necessary to avoid an even greater injustice. Being first virtues of human activities, truth and justice are uncompromising."

Rawls, like other social contract theorists like Hobbes, Locke, Rousseau, and Kant, tries to base the justification of the principle of justice on the agreement of individuals. Metaphysically, they all believe that there is *no* right principles independent of and prior to individuals themselves; instead, justice or right principles must be found in individuals themselves, as Rawls (1980) argues that "the parties to the original

position do not agree on what the moral facts are, as if there were already such facts. It is not that, being situated impartially, they have a clear and undistorted view of a prior and independent moral order. Rather (for constructivism), *there is no such order, and therefore no such facts apart from the procedure as a whole*” (underline added) (p. 568). Therefore, Rawls’ project is to show that his principles of justice is justified not because of any merit or moral desert, but of a social contract; that his principles of justice is not based on any particular conception of the good, but simply on an agreement of individuals.

It should be noted that Rawls’ project is very influenced by Kant’s morality in the sense that both of them try to separate justice as a first principle from any particular conception of the good. For Kant, to act morally is to act as a rational being whose moral duty is determined by the universal maxim rather than his own desires or happiness. This is not to say that desires and happiness are wrong and should be eliminated, but that they are inappropriate to be the *first* principle that every rational being can will as a universal maxim (Kant, 1785). Kant proposes that, as a rational being, we must prioritize the *noumenal* self, whose actions are determined by pure practical reason which is independent from any conception of the good life, over the phenomenal self whose actions are determined by our own conception of the good life. It is not that Kant does not want us to live according to our conception of the good, but Kant thinks that it is wrong to set any particular conception of the good as the moral constraints for all of us, as Kant (1788) argues that “the concept of good and evil is not defined prior to the moral law, to which, it would seem, the former would have to serve as foundation; rather the concept of good and evil must be defined after and by means of the law” (p.65). This idea can be called deontological liberalism (Sandel, 1982).

Rawls (1971), despite the very influence of Kant, tries to make deontological liberalism more acceptable to contemporary political philosophy by insisting that deontological liberalism can be affirmed by the phenomenal self rather than the *noumenal* self (p.264, 587). Since Kant’s project is to put individuals into the *noumenal* world which is very

metaphysically controversial, Rawls tries to justify the moral law without requiring individuals to forgo their own conceptions of the good (Sandel, 1982, p.23). In other words, Rawls' project is to show how individuals who are free to pursue any conception of the good eventually come to agree upon the same first principle. At first glance it seems that Rawls' principles of justice are not separated from a conception of the good, but, indeed, they are separated because, according to Rawls' logic, even though his principles of justice were chosen by individuals who have different conceptions of the good, the fact that they would be accepted by everyone reflects that his principles of justice are not based on any *particular* conception of the good. In other words, since the theories of justice are universally accepted by "every" conception of the good, and not based on any "particular" conception of the good, they are neutral between competing conceptions of the good. In this sense Rawls confirms Kant's deontological liberalism which bases justice as the first principle on the primacy of right rather than any particular conception of the good (Sandel, 1982). Therefore one way to critique Rawls is to show how his theory of justice is actually based on a particular conception of the good rather than some neutrality between them, and this is my primary task in this article.

But it is very important to understand that what Rawls means by 'everyone' here is 'hypothetical' persons, not 'actual' persons; only the former is regarded as a moral agent who is entitled to participate in the deliberation of principles of justice, while the latter is not because they are not yet guaranteed by certain circumstances of justice. In order to come up with a theory of justice, Rawls creates the original position and the veil of ignorance as a thought experiment¹ to reach a social agreement. The original position is the state where everyone is assumed to inhabit *before* entering into the actual society. The veil of ignorance is the idea that people in the original position should be blind to any conception of the good at all in the sense that they must decide any theory of justice without knowing what social position and conception of the good they are going to take until the veil of ignorance is uncovered. In other words,

individuals in the original position must conceive themselves as rational beings who want to maximize their position (the maximin principle), but do not know yet what conception of the good they will pursue. By doing so, everyone would reach the *same* principles of justice which would consist of “(1) Each person is to have an equal right to the most extensive total system of basic liberties compatible with a similar system of liberty for all. (2) Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged, and (b) attached to offices and positions open to all under conditions of fair equality of opportunity” (Swift, 2014, p.25). These two principles of justice determine the roles of the state and political institutions.

His theories of justice have been criticized by communitarians like MacIntyre (1981), Sandel (1982; 1984), Walzer (1983), and Taylor (1985a; 1985b; 1989) that he fails to understand the metaphysical conception of the person; they argue that we could not conceive of ourselves as a rational being whose self is *prior* to our given goods/ends as Rawls assumes in the original position. Thus, Rawls’s principles of justice are not appropriate to do justice for *actual* persons (Mulhall & Swift, 1996). In response to those criticism, Rawls proposes the idea of political liberalism (Rawls, 1977; 1985; 1987; 1988; 1989; 1993; 1999). The main argument of political liberalism is that his two principles of justice do not require that everyone *must* apply liberal doctrine to their personal affairs and think of themselves as a rational being whose self is *antecedent* to their ends/goods *all the time*; they just need to be liberal in political sphere where they need to decide about the best principle of justice for a well-ordered society; his principles of justice do not rule out any ‘reasonable’ moral comprehensive doctrines.

In this sense, his liberalism is different from Kant’s and Mill’s, which are comprehensive liberalism; while Rawls claims that people need to be liberal in political sphere only, not in personal affairs, Kant and Mill claim that they need to be liberal in both of their public and private spheres. As Rawls (1993) argues that “this idea of a shared political life does not invoke Kant’s idea of autonomy, or Mill’s idea of individuality,

as moral values belonging to a comprehensive doctrine. The appeal is rather to the political value of a public life conducted on terms that all reasonable citizens can accept as fair” (p.98), and that “the first difference is that Kant’s doctrine is a comprehensive moral view in which the ideal of autonomy has a regulative role for all of life. This makes it incompatible with the political liberalism of justice as fairness” (p.99). To put it another way, Rawls thinks that it is *possible* to separate political sphere from any moral comprehensive doctrines, and everyone, including communitarian critics, should legitimize his principles of justice as they are just *political*, not *comprehensive*. Next I will show how his principles of justice are not merely political but comprehensive, and thus, his political liberalism cannot claim its neutrality between competing conceptions of the comprehensive doctrines.

It should be noted that what Rawls (1993) really proposes to revise is not about his two principles of justice proposed in Rawls (1971) at all, but he merely revised how we should understand them. As Rawls (1993) writes that “all these elements [two principles of justice] are still in place, as they were in [*A Theory of Justice*]; and so is the basis of the argument for them. Hence I presuppose throughout these lectures the same egalitarian conception of justice as before; and though I mention revisions from time to time, none of them affect this feature of it” (p.7), and what he really wants to propose about political liberalism is that “political liberalism, then, aims for a political conception of justice as a freestanding view. It offers no specific metaphysical or epistemological doctrine beyond what is implied by the political conception itself. As an account of political values, a free standing political conception does not deny there being other values that apply, say, to the personal, the familial, and the associational; nor does it say that political values are separate from, or discontinuous with, other values” (p.10). Therefore, if I can prove that Rawls’ two principles of justice are comprehensive and metaphysical rather than political, then this can prove that Rawls’ political liberalism is also comprehensive, not merely political; Rawls’ political liberalism is inconsistent.

A Theory of the Person and the Impossible Project of Political Liberalism

My main argument is that Rawls has a *particular comprehensive* theory of the person. I intentionally use the term ‘*comprehensive*’ rather than mere ‘*metaphysical*’ or ‘*moral philosophy*’ because although Rawls may claim that his theory is not metaphysical² (Rawls, 1985), he cannot deny that his theory suggests a *particular* conception of the *artificial* person in the original position (Rawls, 1993, p.75). For example, Rawls may say that his political liberalism does not rely on a particular theory of *actual* person, whether they are unencumbered or encumbered selves. This may make him avoid the metaphysical debate about the *real* nature of the person by claiming that his theory does not say that the person described in the original position must be true in the reality. He may reason that he needs to ‘assume’ some certain characteristics³ of the person in the original position just for the sake of representation, and does not mean that he actually thinks that the *real* person must really possess those certain characteristics. But I will argue that this does not make his theory of the person in the original position less comprehensive. The most important question is why Rawls must ‘assume’ those certain characteristics of the person in the original as the way he does? Why does he think that the moral power ‘should’ be given to ‘each individual’ rather than the community?

To criticize Rawls’ political liberalism, I divide the discussion into two main parts: (1) I will show how Rawls’ original position and principles of justice can be criticized from political economy points of view, which will tell us that Rawls’ argument is not based on any social contract, and the secret of his principle of justice lies in a theory of the person; (2) I will show how Rawls’ political liberalism can be criticized from philosophical points of view, which will tell us that Rawls’ argument and his theory of the person is not neutral between competing conceptions of the good as he claims.

A Political Economy Critique of Rawls' Theory of Justice

Before I will discuss this philosophically, I would like to show that even those who agree with him about the role of the original position and that each individual should be given the moral power equally may disagree with him about principles individuals would choose. Harsanyi (1975; 1977) are good examples. He agrees that the original position is an indispensable condition of the social contract, but he disagrees that parties in the original position would always adopt the maximin principle as Rawls claims. Harsanyi (1977) proposes to separate “moral” preferences from “personal” preferences. Moral preferences refer to the decisions of individuals in the situation where they are all ignorant of their actual particular preferences. He believes that to decide this way they all must decide universally as if they take everyone’s preferences into account. Personal preferences refer to the decision of individuals in the situation where they all know their actual places in the society. Even though Harsanyi embraces the same idea as Rawls’ original position, he does not assume that parties in the original position would adopt the maximin principle as Rawls does, and this indicates that different theories of the person (how should they behave?) amount to different characteristics of principles of justice.

Harsanyi (1975) proposes that parties in the original position would choose any principles of justice according to the average utility maximization principle (p.598). He argues that if a society has a mentally retarded individual and other normal individuals, and if the only way to improve the mentally retarded one is an expensive treatment which requires diverting funds from the education of normal people, and if this treatment can only improve the mentally retarded person only slightly, then what principle would parties in the original position choose? According to the maximin principle, they should want the treatment to be done because it improves the least advantaged member of the society. But Harsanyi argues that this is irrational (p.597), and he proposes that parties would adopt the average utility maximization principle which

assumes that everyone has the same probability ($1/n$); they have to choose social systems which yield a higher average utility level. For the sake of simplicity, let's consider the figure 1 as follows.

	P_1	P_2	Average
A	100	200	150
B	145	145	145
C	144	600	372

Figure 1 How would parities in the original position choose?

According to the figure 1, let's assume that there are two persons (P_1 and P_2) whom parties in the original position may turn out to be, and there are three states of the world or social systems (A, B, and C) that might happen to each person. The rows can be read as follows: of the state A, we have a $1/2$ probability of being born as P_1 with 100 utility, and P_2 with 200 utility; of the state B, we have a $1/2$ probability of being born as P_1 with 145 utility, and P_2 with 145 utility; of the state C, we have a $1/2$ probability of being born as P_1 with 144 utility, and P_2 with 600 utility. The question is what social systems or principles of justice we would choose? According to Rawls' maximin principle, we would choose the state B because this social system makes us maximize the least advantaged; since the least advantaged of the state A receives 100, of the state B receives 145, and of the state C receives 144, a rational being who adopts the maximize principle must go for the state or social system B. But, according to Harsanyi's average utility maximization principle, we should choose the state C because this social system makes us maximize the average utility, which is 372, while the state A and B give us only 150 and 142, respectively.

Even if we follow Rawls' experiment of the original position and the veil of ignorance, we do not need to agree with him that his two

principles of justice will always be chosen. This indicates that Rawls' attempt to make the social contract more acceptable has certain problems. Rawls' two principles of justice lies not in the pure procedure of the social contract, but in the certain characteristics of the hypothetical person in the original position. In this sense a theory of the person plays a significant role in determining any principle of justice. Even though Rawls may claim that his theory of the person is merely a thought experiment rather than the metaphysical truth, he cannot deny that the discussion of a theory of the person is very important as different characteristics of the person lead to different principles of justice; thus communitarian critique of Rawls' theory of the person is very relevant and powerful.

Kymlicka (1989) defends Rawls' political liberalism⁴ on the basis that it does not take a "possessive individualist theory of motivation" (pp.887-8). He argues that Rawls's political liberalism does not take the view that "what people want in life is to maximize their share of social resources (rather than promote the good of others), and indeed to maximize their material good (rather than promote their spiritual or emotional well-being)" (p.886). Instead Rawls allows everyone to pursue their own conceptions of the good life; no one is barred from taking other people's welfare and spiritual well-being into account. This implies that Rawls' political liberalism is far from being *comprehensive* liberalism which biased against some *particular* way of life. To put it simply, according to Kymlicka and Rawls, political liberalism does not take any particular comprehensive theory of the person as everyone is still allowed to take any comprehensive theory of the person as they see fit. But I believe this argument is flawed.

First of all, Kymlicka and even Rawls himself misunderstand that communitarians like MacIntyre, Sandel, and Taylor discuss a theory of the person "before" the original position and the veil of ignorance are taken out, not "after" that. As I just said, a theory of the person in the original position is very important to justify the principles of justice because different theories lead to different principles, and we have seen that Rawls' theory of the person is merely one of possible alternatives

rather than the only conceivable one, so whatever theory of the person Rawls chooses, he cannot deny that it is based on a particular conception of the person rather than a neutrality between competing conceptions. Before I will continue discussing a theory of the person, I would like to show that, indeed, Rawls' social contract theory is not really a social contract argument because according to any social contract theory, a just principle must be justified on the basis of individual agreements rather than of any independent principle.

For example, if P_1 and P_2 come to agree upon the social system B, then we say that the state B is just not because it is just by itself but because it is the result of the contract, and if both P_1 and P_2 change the content of the contract to other systems (A or C), then we say that the state B is not just anymore. This means that any just principle of justice must be the result of the contract, not something just in itself. In this sense Rawls' argument is far from being called the social contract because his argument does not allow any changes of the content of the contract at all, and, most importantly, his theory of the person is characterized by himself "before" any actual person and contract occur; this means that his theory of the person is justified independently of any result of the contract. As I said, since different theories of the person lead to different principles of justice, and Rawls' theory of the person is assumed "before" any social contract happens, Rawls cannot deny that his principles of justice are justified not based on the social contract, but on his particular conception of the good. I will defend my argument through the discussion of Hampton (1980)'s argument.

For Hampton, any contract must be the procedure of mediation between two or more parties who have *different* preferences and needs⁵, and the contract must be irrevocable until every party voluntarily agrees to end or redo the contract; any contract must be based on the *procedure* rather than the *finality* (the content of the contract). For example, if Mr.A lends Mr.B a certain amount of money and Mr.B promises to repay Mr.A that certain amount of money with another certain amount of interest, the contract must be irrevocable *until* both Mr.A and Mr.B voluntarily

agree to end or redo some characteristics of the contract. This is the *reciprocity* condition. But in the original position *every* party is forced to have the *same* preference (risk-aversion) and goal (maximin rule), to face the *same* situation (complete uncertainty); everyone is forced to be the same person and they actually contract with themselves rather than other people at all. Even though Rawls may argue that they need to take into consideration other people's needs (the demands of reciprocity), he fails to understand that the reciprocity condition requires two things that are unavailable in his original position: (1) every party must be allowed to represent his own preference and advantages from their point of view; (2) every involved party must be allowed to end/redo the agreement at any time they all want. All of this means that the reciprocity condition cannot ensure any principle as the finality because it just ensures that each party voluntarily agrees upon an agreement, whatever characteristics of the contract. But Rawls seems to suggest that the original position, plus the veil of ignorance, require the reciprocity condition, and thus the two principles of justice are chosen by a social choice process.

Hampton rightly argues that Rawls' two principles of justice are actually chosen by *individual* choice, not *social* choice because the reciprocity condition cannot guarantee any principle as the *unconditional* finality. Thus even though every party in the original position may actually choose the two principles of justice, they are entitled to change them at any time they *all* agree to do so; the two principles of justice must be contingent, conditional, and revocable. As we have seen that parties in the original position do not necessarily need to adopt the maximin principle and choose two principles of justice as Rawls expects, therefore to prevent any other principle than the two principles to be reached, Rawls needs to 'force' every party to accept the two principles as the *unconditional* finality. Now we can see that Rawlsian theory is inconsistent; if he claims that any principle of justice is justified *only if* it is the result of the contract, then he must not take any contract as the unconditional finality, but if he claims that the two principles are *always* the finality, then he must rule out any attempt to redo the contract even though every party agrees to do so.

For Hampton, Rawlsian contract is merely an individual choice under conditions of uncertainty; an individual has to choose the best possible outcome for himself without a consideration of other people because there are no longer other people; even though there are a million people, there are no real “other people” in the Rawlsian contract insofar as everyone is forced to have the same preference (the more, the better) and same way of reasoning (maximin rule). According to Rawls, if one is not sure if he will be born in a rich or poor family, then he needs a principle of justice that makes him better off if he actually turns out to be a poor, and since everyone thinks in the same way, everyone finally agrees upon the same principle of justice. This seems to be the result of contract, but actually it is not. The key word lies in the assumption that ‘*everyone is the same*’ which implies that everyone has *no* different goals; everyone has the same goal in the first place. The question is that if there is no any different goal among parties, then on what point do they need any contract in the first place? If there is no conflict about the way of reasoning and they are all the same person, why do they need to come to contract with each other in the first place? In this sense Hampton is right in saying that there is *no* contract in the original position because everyone has the same preference and goal, thus there is no need to have a contract among them in the first place.

Now I want to show how, according to Hampton’s argument, Rawls’ original position does not raise problems of social choice at all. I will do so by using my own representation. The real problem (puzzle) of social choice is about how to reach an agreement among *different* preferences. Problems of social choice are raised in Arrow’s Impossibility (Arrow, 1951). One Impossibility is the ‘*unrestricted domain*’ axiom which requires that *all* preferences are allowed, but Rawlsian theory violates the unrestricted domain as it rules out some certain preferences in the first place. Let me make this argument more concretely. Let’s assume that there are three persons – Rich, Middle, Poor – in which Rich is assumed to be the most advantaged, Middle is the middle advantaged, and Poor is the least advantaged, and there are three policies (A, B, C). Let’s assume

also that Policy A is most favored by Rich, B by Middle, and C by Poor. This can be put in table as the following figure 2:

	Rich	Middle	Poor
Most Preferable	A	B	C
Middle Preferable	B	C	A
Least Preferable	C	A	B

Figure 2 Problem of Social Choice⁶

From the Figure 2, any social decision cannot be made as the following reasons. According to Rich, $A > B > C$, hence we can say that $A > B$ and $B > C$. According to Middle, $B > C > A$, hence we can say that $A < B$ and $B > C$. According to Poor, $C > A > B$, hence we can say that $A > B$ and $B < C$. We can see that Policy A is preferred over Policy B by a 2-1 margin, that is, Rich and Poor prefer A to B, while Middle prefers B to A; and that Policy B is also preferred over Policy C by a 2-1 margin, that is, Rich and Middle prefer B to C, while Poor prefers C to B. Since there are three persons in the society, if we use the majoritarian rule, then the society should choose Policy A because A is preferred over B, and B over C. However, this social choice does not hold because of its irrationality. According to the principle of rationality, individuals must be able to tell what they prefer (completeness) and the correct order of their preferable choices (transitivity). For example, someone is rational if he can tell that he prefers A to B, and B to C (complete), and he can also tell that he prefers A to C (transitive) (Hausman & McPherson, 1996). Therefore, in the figure 2, if A is preferred over B and B over C, then A must be preferred over C always. But this is not the case as, indeed, we found that C is preferred over A, not vice versa; that is, Rich is the only one who prefers A to C, while Middle and Poor prefers C to A. Therefore, from the figure 2, we can have the following social ordering: $A > B$ and $B > C$, but $A < C$, which is irrational and any social choice or principle

of justice cannot be made through the social contract.

According to Arrow's Impossibility Theorem, insofar as those ranking preferences from the figure 2 are allowed, any social decision cannot be made because any social rank is inconsistent. According to Sen's Liberal Paradox, any social decision cannot be made because any social decision violates the minimal liberty (Sen, 1970; 1983; 1984). This raises problems of social choice: how to make a social decision which is transitive (Arrow) or consistent with the minimal liberty (Sen). But the Rawlsian contract would solve this problem by 'forcing' three of them to be the *same* person who faces the same situation, that is, no one actually knows their own preferences and that of others. In the original position, for example, Rich cannot claim his own preferences ($A > B > C$) in the first place because he may turn out to be Middle and Poor equally. Therefore, he must think of the worst case (Poor) and, according to the maximum rule, he must choose a policy that maximizes his worst outcome, that is, he must choose Policy C. At the end, everyone should prefer Policy C, and this seems to be a social decision. But we can see that this is not because it violates the unrestricted domain axiom. According to Hampton (1980), the original position is the place where there are *no* differing preferences and all people are "faced with the same problem of uncertainty" and "there is only a series of independent but identical deliberation" (p. 326). Therefore, Rawlsian contract does not actually raise problems of social choice, and cannot be called a social contract.

The question is if Rawls' two principles are not justified by a social contract (voluntary agreement), then what justifies them? I suggest that the original position and the veil of ignorance are not just conditions of justice as Rawls claims, but they are the normative justification of the two principles itself. The original position already provides what is justified and what is not by ruling out some certain motivations and preferences in the first place (restricted domain). It is not surprising at all that individuals in the original position choose the two principles because they are 'forced' to do so in the first place. Rawls' argument is at best a rational choice theory, not a social choice theory: while the former assumes a certain

motivation, the latter does not. Now we can see that Rawls' justification of his two principles of justice cannot be justified by social contract argument, but by his own particular conception of the good which is realized in the form of the circumstances of justice; since different circumstances of justice and theories of the person lead to different principles of justice, Rawls' secret of two principles of justice lies not in the social choice, but in his particular theory of the person. The next question I will deal with is if Rawls' theory of the person is comprehensive and metaphysical?

A Philosophical Critique of Rawls' Political Liberalism

I think that political liberalism is comprehensive because it is very *individualistic* and biased against some other ways of life. This is absolutely not to say that political liberalism and individualism as such are morally wrong, but simply that the *fundamental* idea of political liberalism, which is the belief that it is free of any comprehensive and metaphysical controversies and that it has no says about private sphere, is indefensible. Even though it is true that it does not say that the *only* motivation of individuals is to maximize their own utility or be self-interested, it obviously says that each individual *equally* has the moral power to choose their own conceptions of the good life. The question is that if Rawls claims that he does not take any particular comprehensive theory of the person, why does he assume that it is 'each individual' who should have the moral power rather than the community or why everyone should be morally equal in the first place? According to Sandel (1982; 1984), human beings are encumbered selves in the sense that our ends are *prior* to our selves and our community is constituted of our selves, so it is wrong to conceive a person as an unencumbered self⁷ whose self is prior to the good and the community is conceived in the sentimental sense.

Rawls would argue that political liberalism does not demand that everyone must conceive of themselves as an unencumbered self; instead, insofar as they 'voluntarily' choose to be encumbered selves, no one can intervene in their decision. But this argument fails to realize that it still

gives the moral power to ‘each individual’ in the first place. Even though people may have different kinds of motivation and conceptions of the self, their actions are still justified/legitimated on the basis of ‘freedom of choice’ of individuals, and an encumbered self is important at best as the second-order justification. As we have seen and will see more, Rawls’ principles of justice, which were chosen by rational/unencumbered beings in the original position, are the first virtue of all institutions and even our private lives in such a way that if there is a conflict between principles of justice and our subjective conceptions of the good, the former must be decisive always. In this sense Rawls always gives the priority of principles endorsed by an unencumbered self over any other principles. In other words, Rawls conceives the person as an unencumbered self all the time. To understand my argument clearer, I will show how Rawls takes a *particular* comprehensive theory of the person in the original position and why his difference principle is very comprehensive, and then why he fails to distinguish between political sphere and comprehensive sphere.

In the original position, Rawls takes the deontological theory of the person in the sense that everyone is assumed to be blind to any conception of the good, and they must think of themselves as a self-interested unencumbered self. Although Rawls may claim that this is just a device, he cannot deny that the characteristic he assigns to the person is merely one of all possible alternatives, as I have shown in the previous section, and more importantly, he cannot deny that he *already* accepts the possibility of the self-interested motivation. Why does he think that one ‘can’ be self-interested? Of course, he does not say that one ‘should’ be self-interested, but he actually thinks that it is ‘permissible’ and ‘acceptable’ for one to be self-interested. He should be aware that to be called ‘moral comprehensive doctrines’ (non-neutrality), one does *not* need to suggest that the right thing to do is to do a *particular* action and think in a *certain* way, but, indeed, it is enough to judge that what action is permissible and what action is not. In this sense, political liberalism, which insists that one ‘can’ do whatever they want, given that they do not *physically*⁸ harm other people, is a kind of comprehensive doctrine despite the fact

that it does not suggest what *particular* actions people ‘should’ do. For example, liberals may say that one ‘can’ choose whether to donate to a charity or not. Of course, they do not say that one ‘should’ or must donate to a charity, but they accept the possibility of both choices (donate and not donate) in the first place. This means that they already judge that it is acceptable *not* to donate, so their judgement is based on the premise that one should be conceived as an unencumbered self.

I argue that there are two main answers to the question of why we should hold that everyone is morally equal in the first place? One is metaphysical/comprehensive, another is political. I will show that only the former can be defensible, while the latter, which is held by Rawls, cannot.

A Comprehensive Theory of the Person

There are many reasons for accepting the notion that everyone should be morally equal. John Locke (1690) argues that everyone has natural rights to life, liberty, and property which are determined by natural laws (p.9). It should be noted that these natural rights and laws are given by God. As Locke (1690) writes that,

In transgressing the law of nature, the offender declares himself to live by another rule than that of reason and common equality, which is that measure God has set to the actions of men, for their mutual security; and so he becomes dangerous to mankind, the tie, which is to secure them from injury and violence, being slightly and broken by him (p.10).

God, who hath given the world to men in common, hath also given them reason to make use of it to the best advantage of life, and convenience (p.18).

As much as any one can make use of to any advantage of life before it spoils, so much he may by his labour fix a property in: whatever is beyond this, is more than his share, and belongs to others. Nothing was made by God for man to spoil or destroy (pp.20-1).

God having made man such a creature, that in his own judgment, it was not good for him to be alone, put him under strong obligations of necessity, convenience, and inclination to drive him into society, as well as fitted him with understanding and language to continue and enjoy it (p.42).

God and nature never allowing man so to abandon himself, as to neglect his own preservation: and since he cannot take away his own life, neither can he give another power to take it (p.88).

Locke's argument is based on a very comprehensive/metaphysical theory of the person as he argues that a person's essence (true nature) is given by God, and to act morally is to act according to God's command. Many would find the notion that everyone should be morally equal because it is the command of God is very comprehensive and vulnerable to objection as the existence of God, is extremely controversial. Even though I agree that Locke's argument is quite controversial, I do not think that we can defeat his argument by claiming that there are better arguments which are less controversial/comprehensive. In other words, I would like to suggest that if we want to defeat Locke's argument, we cannot argue that his argument is wrong because it is controversial, but because it is too unreasonable; since every theory of the person is very controversial and comprehensive, so the best argument is not the one which is less controversial/comprehensive, but the one which is more reasonable. This is a very important issue which I will discuss later.

Aristotle argues that a person's essence is a social being whose ends are given by its own teleology or merits, which he argues that happiness

is the most noble virtue we all should attain in our lives together, and governments should exist not only to protect freedom but also to promote those virtues/goods of individuals. Aristotle's theory of the person is different from that of Locke in that the *telos* or nature of a person is not given by God, but by the nature itself, as Brown (2009) argues that "there is a way human beings ought to be and ought to live. This is not because god created them for a purpose – something Aristotle did not hold – but simply because they are a certain kind of living being, and every living species has its own work or function" (p. xi). For Aristotle, to act morally is to act according to the nature of the self, which is to live a happy life, as he writes in *the Nicomachean Ethics* that:

Happiness, then, is something final and self-sufficient, and is the end of action. (Book I, 1097b)

If this is so, virtuous actions must be in themselves pleasant. But they are also good and noble, and have each of these attributes in the highest degree, since the good man judges well about these attributes; his judgement is such as we have described. Happiness then is the best, noblest, and most pleasant thing in the world (Book I, 1099a).

Now virtuous actions are noble and done for the sake of the noble. Therefore the liberal man, like other virtuous men, will give for the sake of the noble, and rightly; for he will give to the right people, the right amounts, and at the right time, with all the other qualifications that accompany right giving; and that too with pleasure or without pain; for that which is virtuous is pleasant or free from pain – least of all will it be painful (Book IV, 1120a).

It is not the capacity that makes the boaster, but the purpose; for it is in virtue of his state of character and by being a man of a certain kind that he is a boaster (Book IV, 1127b).

We now can see that, for Aristotle, to act morally is to act not according to our subjective pleasures or happiness but according to the nature of the things themselves. The right thing to do exists prior to individuals and individuals must discover the *telos* or virtue of any social practice. For example, they should realize what characteristics are required for one to be called a good person. Moreover, Aristotle also believes that the purpose of government is to help men pursue noble actions, as he writes in *the Politics* that:

It is clear therefore that the state is not an association of people dwelling in the same place, established to prevent its members from committing injustice against each other, and to promote transactions. Certainly all these features must be present if there is to be a state; but even the presence of every one of them does not make a state *ipso facto*. The state is an association intended to enable its members, in their households and the kinships, to live *well*; its purpose is a perfect and self-sufficient life ... So we must lay it down that the association which is a state exists not for the purpose of living together but for the sake of noble actions (Book III, 1280b29-1281a2).

Now we can see how Locke and Aristotle reason for their comprehensive/metaphysical theories of the person. For Locke, a person is created by God, and doing the right thing is to act according to God's command; for Aristotle, a person is created and defined by the nature/*telos* of itself, and doing the right thing is to act according to the highest nature of the self. So far, I may be criticized for failing to acknowledge that both Locke and Aristotle do not hold that everyone is morally equal as some may be masters some may be slaves. I would argue that my real point of discussing them is not to say that both of them really give moral power to each individual equally, but that their arguments can support the notion metaphysically. In other words, I suggest that one may borrow this argument (God's command – Locke; Teleology – Aristotle) to advocate

the notion that everyone should be morally equal.

Now let me explore another comprehensive/metaphysical theory of the person of Immanuel Kant. According to Kant (1785), a person is perceived to live in two different worlds, the *noumenal* and empirical/phenomenal world⁹; but to be a moral person, one needs to move from the empirical world, in which his actions are determined by their own different empirical conditions and conceptions of the good, to the *noumenal* world, in which his actions are determined by pure practical reason or universal moral laws; only in the *noumenal* world that a person can realize his real self or higher nature¹⁰ of the self, which is a rational being. As Kant (1785) writes that:

Hence he has two points of view from which he can regard himself, and recognize laws of the exercise of his faculties, and consequently of all his actions: *first*, so far as he belongs to the world of sense, he finds himself subject to laws of nature (heteronomy); *secondly*, as belonging to the intelligible world, under laws which being independent on nature have their foundation not in experience but in reason alone (pp.84-5).

If therefore I were only a member of the world of understanding, then all my actions would perfectly conform to the principle of autonomy of the pure will; if I were only a part of the world of sense they would necessarily be assumed to conform wholly to the natural law of desires and inclinations, in other words, to the heteronomy of nature. (The former would rest on morality as the supreme principle, the latter on happiness.) (p.86).

What he morally “ought” is then what he necessarily “would” as a member of the world of the understanding, and is conceived by him as an “ought” only inasmuch as he likewise considers himself as a member of the world of sense (pp. 87-8).

In this sense Sandel (2009) concisely articulates Kant's theory of the person: "Kant argues that every person is worthy of respect, not because we own ourselves but because we are rational beings, capable of reason; we are also autonomous beings, capable of acting and choosing freely. Kant doesn't mean that we always succeed in acting rationally, or in choosing autonomously. Sometimes we do and sometimes we don't. He means only that we have the capacity for reason, and for freedom, and that this capacity is common to human beings as such" (p.105). Therefore, according to Kant, the fact that a human being is capable of reason is the foundation of the moral action. One may argue that Kant's theory of the person is more acceptable than that of Locke and Aristotle because it does not base moral judgement on God and the essence of the person, which is quite controversial, but I would argue that Kant's theory is not less controversial than that of Locke and Aristotle.

It should be noted that there are clearly two different statements of the notion "since we are all rational beings who are capable of reason, so we should act according to the moral laws determined by rational beings"; the first statement, which is "we *are* all rational beings who are capable of reason," can be scientifically proved; the second statement, which is "we *should* act according to the moral laws determined by rational beings," cannot be scientifically proved because it is a normative statement engaging with the language of moral judgement, e.g. good or bad, moral or immoral, and so on. In this sense moral and political philosophy which must engage in judging what's the right thing to do is scientifically unprovable; it may be provable that doing action A may make me happy, but it is unprovable that I should do action A, as it may be the case that I should do something else despite it does not make me happy; moral judgement does not lie in the scientific/provable facts, but in a given unprovable premise.

In this sense I may accept that "we are all rational beings who are capable of reason" as it can be scientifically proved, and this is less controversial; but despite of that acceptance, I do not need to follow that "so we *should* act according to the moral laws determined by rational

beings.” Why do we use the sentence “Since ..., So ...” here? I would argue that these two sentences can be combined rationally *only if* we already give our unprovable moral values to the first sentence; that is, we must already value the fact of the capacity for reason in the first place, otherwise one can refuse to accept the second sentence that “so, we should act according to the moral laws.” I agree with Taylor (1985b) who brilliantly argues that “the mere possession of the capacity will have no normative consequences at all for us if we do not share the conviction that this capacity commands respect” (p.192), and that “beings with these capacities command our respect, because these capacities are of special significance for us; they have a special moral status” (p.193). In this sense Kant’s theory of the person is also controversial as it is based on an *unprovable* moral value he gives to the world in the first place; he must see that the capacity for reason is good in itself, otherwise the mere fact that we are all capable of reason cannot support his moral viewpoint.

It is worth noting that my above argument can be applied to other similar notions, e.g. “since individuals are the smallest atomistic creatures, so they should have equal moral power.” Again, this notion is controversial and based on a particular comprehensive/metaphysical theory of the person, as those who hold this notion must value the being of the smallest atomistic creature in the first place, otherwise being of this creature would be nothing to with its normative statement at all. So far, we have seen that a theory of the person is indispensable because different theories mean different principles of justice, and that any theory of the person must be based on any particular comprehensive and metaphysical doctrines. The next question is if Rawls’ attempt to create a theory of the person rather than comprehensive and metaphysical is defensible, that is, is his political liberalism defensible?

A Political Theory of the Person: Why Is Rawls' Political Liberalism Indefensible?

Having seen that many previous theories of the person must engage in controversial comprehensive doctrines in one way or another, Rawls' project is to construct a theory of justice without relying on any comprehensive theory of the person in order to make his principle more acceptable as it does not involve any controversies about the nature of the person; as Rawls (1977, p.165) argues that

The Problem is this: to develop a viable Kantian conception of justice *the force and content of Kant's doctrine must be detached from its background in transcendental idealism and given a procedural interpretation by means of the construction of the original position.* It must then be shown that the construction, which draws upon the idea of the social contract, is *not subject to the cogent objections* that idealists raised against the contract doctrine of their day ... I have tried to show how *the conception of justice as fairness avoids these and similar objections.* The procedural interpretation of Kant's view not only satisfies *the canon of a reasonable empiricism,* but its use of the idea of the social contract meets Hegel's criticisms [italics added].

Rawls' political liberalism aims to complete two tasks; one is to preserve a Kantian conception of justice, another is to replace Kant's transcendental idealism, which is comprehensive, with a reasonable empiricism, which Rawls thinks is not comprehensive as it is based on the idea of the social contract. If he can accomplish these two tasks, then Rawls' political liberalism is defensible. However, I argue that Rawls' political liberalism is indefensible because these two tasks are theoretically impossible in the first place. As I have shown, Rawls' argument is not based on the idea of the social contract, but on his particular theory of the person; since different theories of the person lead to different principles of justice. Rawls (1977) himself accepts this fact: "once we think of the members of society as free and equal moral persons, it is

natural for a social contract view to take the basic structure of society as the primary subject of justice” (p.164). Rawls’ principles of justice are impossible without his particular theory of the person, which is a person who is morally free and equal. I will argue in this section that, however reasonable his theory of the person is, his theory of the person is also comprehensive/metaphysical. This is not to say that everyone should be morally free and equal, but that the notion is based on a particular comprehensive conception of the person/good.

Rawls would argue that his theory of the person is merely a device of representation is untenable because, as I have shown in the section of “*A Political Economy Critique*,” his theory is just one of all possible alternatives; even though I agree with him that a theory of the person is simply a device of representation, I do not need to agree with him that a person should be blind to any conception of the good or morally equal and free, and so on; therefore Rawls’ theory of the person must be based on a particular comprehensive conception of the person. To defend his political view, Rawls can criticize my argument with two possible arguments; first he may reply that his theory of the person is political, not comprehensive because it is widely accepted by democratic values; second he may reply that this is because it is more reasonable than other theories of the person, that is, it is simply a “thin” theory of the person, which assumes simply a quite general characteristics of the person (Rawls, 1971).

Let me discuss the first reply. Rawls (1999) argues that

I suggest that *we leave aside how people’s comprehensive doctrines connect with the content of the political conception of justice* and, instead, regard that content as arising from the various fundamental ideas drawn from *the public political culture of a democratic society*. Putting people’s comprehensive doctrines behind the veil of ignorance enables us to find a political conception of justice that can be the focus of an overlapping consensus and thereby serve as a public basis of justification in a society marked by *the fact of reasonable pluralism* [italics added] (p.32).

This is *the fact that citizens in a pluralist liberal democratic society realize that they cannot reach agreement, or even approach mutual understanding, on the basis of their irreconcilable comprehensive doctrines*. Thus, when citizens are discussing fundamental political questions, *they appeal not to those doctrines, but to a reasonable family of political conceptions of right and justice*, and so to the idea of the politically reasonable addressed to citizens as citizens. This does not mean that doctrines of faith or nonreligious (secular) doctrines cannot be introduced into political discussion, but rather that citizens introducing them should also provide sufficient grounds in public reason for the political policies that religious or nonreligious doctrines support [italics added] (p.125).

Therefore this reply/argument can be articulated as follows: since a democratic society requires pluralism which enables its members to pursue their own comprehensive doctrines in the private sphere, it is unreasonable to impose any particular comprehensive doctrines on the whole society as other members may not share the same doctrine; it is more reasonable to use a political principle like democratic value or pluralism to govern the whole society; it does not embrace any particular comprehensive doctrine, but simply a political value. In other words, persons should be conceived morally free and equal not because it is metaphysically true but because it is reasonable to a democratic society. Despite the soundness of this argument, it is far from being neutral among any particular comprehensive doctrines.

It may be true that to preserve a democratic society, we need political liberalism, political conception of the person, reasonable pluralism, the public political culture of a democratic society, and so on, but the question is why a democratic society *in the first place*? Why should a democratic society be valued *in the first place*? Rawls may argue that this is because a democratic society is best suitable for individuals who are morally free and equal, but this indicates that a theory of the person is the foundation of his argument. Thus the real question would be why *should* we hold that

every individual is morally equal and free *in the first place*? This question can be answered straightforwardly through comprehensive doctrines, e.g. this *should* be so because this *is* a metaphysical truth, and doing the right thing is to fulfill the nature of the person, and so on. But this answer is clearly impossible for political liberalism; instead, it must be answered through political conceptions, e.g. this should be so because it is reasonable for individuals who are morally free and equal. And the old question would be raised again: why should we hold that individuals are morally equal free and equal in the first place? At the end, political liberalism would become merely an *arbitrary* principle which lacks foundation.

Now let me discuss the second reply. As I said, the secret of Rawls' principles of justice relies on his theory of the person in the original position, which is articulated as the "circumstances of justice," so I will discuss why Rawls' circumstances of justice are comprehensive. It should be noted that Rawls himself accepts that he needs some conception of the good to form the person in the original position, but this conception must be too thin and weak to be called a comprehensive doctrine, as Rawls (1971) argues that,

But to establish these principles it is necessary to rely on *some notion of goodness*, for we need assumptions about the parties motives in the original position. Since these assumptions must not jeopardize the prior place of the concept of right, the theory of the good used in arguing for the principles of justice is restricted to the bare essentials. This account of the good I call *the thin theory: its purpose is to secure the premises about primary goods requires to arrive at the principles of justice*. Once this theory is worked out and the primary goods accounted for, we are free to use the principles of justice in the further development of what I shall call the full theory of the good [italics added] (p.396).

There is no inconsistency, then, in supposing that once the veil of ignorance is removed, the parties find that they have ties of sentiment and affection, and want to advance the interests of others and to see their ends attained. But the postulate of mutual disinterest in the original position is made to insure that *the principles of justice do not depend upon strong assumptions*. Recall that the original position is meant to incorporate *widely shared and yet weak conditions*. A conception of justice should not presuppose, then, extensive ties of natural sentiment. At the basis of the theory, one tries to assume as little as possible [underline added] (p.129).

But a human society is characterized by the circumstances of justice. The account of these conditions *involves no particular theory of human motivation*. Rather, its aim is to include in the description of the original position the relations of individuals to one another which set the stage for questions of justice [*italics added*] (pp.129-30).

They assume that *they would prefer more primary social goods rather than less*. Of course, it may turn out, once the veil of ignorance is removed, that some of them for religious or other reasons may not, in fact, want more of these goods. But from the standpoint of the original position, *it is rational for the parties to suppose that they do want a larger share*, since in any case they are not compelled to accept more if they do not wish to, nor does a person suffer from a greater liberty [*italics added*] (p.142).

We now can see that Rawls distinguishes between the *thin/weak* theory of the good/person and the *full/strong* theory of the good/person; while the former is acceptable to a party in the original position, the latter is not. Please notice that, despite his use of conceptions of the good, he still believes that his conception of the good is not based on a particular comprehensive theory, as Rawls (1993) argues that “the veil of ignorance, to mention one prominent feature of that position, has no specific

metaphysical implications concerning the nature of the self” (p.27). Rawls believes that by using the *thin/weak* theory of the good/person, which is the “*the more, the better*” motivation and the maximin motivation, he can avoid any comprehensive doctrines. But I will argue that his thin/weak theory of the good/person is also comprehensive.

First of all, how do we distinguish between the thin and full theory of the good? What is the definition of the thin theory? Rawls does not discuss these issues in detail, but we can articulate as follows: the thin theory is any characteristics of the person which are widely accepted and less controversial; the full theory is any characteristics which are less common and quite controversial. But this definition is controversial itself, as Sandel (1982) asks that “weak or strong with respect to *what?*” (p.45). I am sure that Rawls would say that a self-interested motivation is a weak/thin theory, while benevolence and altruism are full and strong. The question is how does he know that this is true? Rawls can only answer this question in two different ways.

First, he might argue that this is true because of empirical facts; since many or most people see it that way, he also sees it that way, too. But this is problematic. While it is true that many modern societies would accept that thin-full distinction, it is also true that many pre-modern societies may not; from those pre-modern societies’ points of view, benevolence/altruism may be more widely stressed than self-interest, and so it should be a thin theory to them. Rawls might argue that we do not need to take those pre-modern societies into account because they are irrational. But this raises another serious problem: he must judge who is rational/irrational by his own particular conception of the good because he must assert this even before any social contract begins. Let me assume that there are only 100 people in the world, and that 90 of them hold that self-interest is a thin theory while benevolence is a full theory, and another 10 hold otherwise. In this sense it is clearly that the notion that self-interest is a thin theory is more widely accepted, and so Rawls should hold this view. But if he really holds this view, then what is the difference between his liberalism and his enemies, utilitarianism, which

favors the majority's views? Don't forget that Rawls' theory of justice is constructed to overcome utilitarianism in the first place, so Rawls cannot hold this view.

Second, he might argue that this is true because of its rationality itself; since people may have different views concerning the thin and full conceptions, we must assert that it is always rational to assume that self-interest is thin, and benevolence is full. But this reply is also problematic: how do we know if a theory is more or less controversial and widely accepted if we do not refer to actual persons? Since the concept of 'controversy' would be meaningless if it is beyond real persons' opinions, Rawls now has only two choices. First, he must embrace only the first reply, which holds that the distinction between the thin and full theories depends on empirical facts. Second, he must embrace only the second reply, which holds that the distinction between the thin and full theories depends on his own conception of rationality. Whatever he chooses, his political liberalism would collapse because for the first choice, empirical facts cannot support his theory, and for the second choice, his theory is based on a comprehensive/controversial doctrines.

I also want to argue further that even though he is successful in distinguishing between the thin and full theories of the good, his thin theory is also comprehensive. Now let's assume that his definition of the thin theory is sound and clear, the question is why should we hold that the thin theory is more reasonable than the full theory? As I have argued, Rawls cannot refer to the majority's views as it would undermine his whole theory. The only reason is that he already values the thin theory as such, and this will also undermine his political liberalism as the following reason. For instance, if a person really believes in God's command which teaches him that everyone is not morally free and equal, and everyone is not born by accident but by God's intention; it is not accident that someone is born rich and someone is born poor, and so on, thus any income redistribution is unjust because it violates God's command. How would Rawls respond? I guess he would say that his belief is too strong to be accepted by everyone, so his belief cannot become the theory of justice.

But, as we have seen, it is impossible for any single principle, *weak or strong*, to be accepted by everyone. Rawls would argue that he means only every ‘rational’ person, not everyone in the usual sense. But where does his conception of rationality come from? It is clearly not from actual persons’ opinions, but from his own reasoning.

One may argue that we should not give the moral power to those who believe that everyone is not morally equal and free because they violate the rights in the first place. But this argument is clearly irrelevant and misplaced because we must realize that Rawls’ theory of the person is formed “before” any principle of justice occurs, and, as Rawls claims, “before” principles of justice come to play no one can claim his “rights” absolutely. But Rawls (1971) seems to violate his own theory by asserting that “these assumptions [conception of goodness] must not jeopardize the prior place of the concept of right” (p.396). This means that, for Rawls, what is just/unjust can be judged “before” any principle of justice is chosen, as Rawls (1971) himself accepts that “the idea of the original position is to set up a *fair procedure* so that any principles agreed to will be just” [italics added] (p.136), and that “if anyone after *due reflection* prefers a conception of justice to another, then they all do, and a unanimous agreement can be reached” [italics added] (p. 139), and that “if a knowledge of particulars is allowed, then the outcome is biased by arbitrary contingencies” (p.141).

Let’s us discuss these passages critically. Indeed, the ultimate moral judgement of his theory of justice is not based on the agreement of the social contract, but on the circumstances of justice (*fair procedure/ due reflection*) characterized by Rawls himself. As he accepted, simply the fact that people come to agree upon the contract cannot guarantee that the really just principles of justice would be reached, this is why he must propose that the agreements would be just *only if* they are proceeded *under* a fair procedure characterized by “his” certain characteristics. The notion that particular knowledge is arbitrary contingency may seem rational to many people, but this clearly cannot be the case for everyone; thus his notion is also controversial and comprehensive as it is merely a

particular theory. Since Rawls' principles of justice are already guaranteed by his *own* circumstances of justice, which holds that it is "*unjust*" for a person to know his particular conceptions of the good before the contract, the ultimate justification of his theory is based on his own particular comprehensive doctrines rather than the social contract; that is, he already knows what is just and unjust *before* any social contract happens.

According to the difference principle, Rawls embraces two different conceptions of the *comprehensive* doctrine: one is a *particular* assumption of the human motivation, another is the *deontological* theory of the person. For the first conception, it would be impossible to talk about economic (dis)advantages if we could not make a comparison between each individual's good/end. He needs to assume that 'everyone' prefers 'more to less' and material wealth is very important to them. Again, this assumption seems very common to many of us, but there is no good reason to believe that this is what 'everybody' *actually* wants unconditionally; thus, to make the difference principle possible he needs to lean on a *particular* assumption of the human motivation which is not neutral/political.

For the second conception, the difference principle reflects his deontological theory of the person in a deep sense. It is true that he allows 'actual' persons to pursue their own goods and properties as they see fit, but it is also true that he does not allow 'anyone' to be entitled to their own goods and properties *absolutely*; instead they all have to be subject to the difference principle which emerges from the agreement among *deontological* persons. All of this means that his theory of justice is more acceptable to those whose comprehensive doctrine is *individualism* than those whose comprehensive doctrine is not. Let's us come back to my example of someone who believes that any income redistribution is unjust because it violates God's command. Rawls would argue that those claims are too 'strong' and 'contingent' to be a 'general' conception of the person who is rational enough to choose the best principles of justice. But this is problematic; since he finds that some 'actual' persons are irrational and are not suitable to do the task of discovering the principle of justice, the first task would be to make them 'rational' by forcing them to conceive

themselves as an unencumbered self. In other words, Rawls has to insist that those who have such ‘strong’ conceptions of the person are *irrational*, and they need to ‘change’ their minds by accepting the notion that we are all born *accidentally*, not according to our choices in previous lives, or by God’s command. This is not to say that it is reasonable, but that Rawls cannot avoid choosing a *particular* conception of the good/person which is very comprehensive.

Rawls may argue that this conception is not comprehensive/metaphysical because he still allows anyone to hold that we are all born as a result of our choices in previous lives, by God’s intention, but only if he or she accepts the difference principle. But this is very paradoxical since the foundation of the difference principle is the belief that we are all born by accident, thus any belief that holds otherwise can be accepted at best as the second-order truth. Since Rawls’ two principles of justice are always the first virtues of the whole society, to accept them is to also believe that his theory of the person in the original position is metaphysically true. Rawls may argue that his theory of the person is simply a device to represent his philosophy, and actual persons do not need to hold it in the reality. But if this is really the case, then why doesn’t he allow actual persons to hold their “strong” conceptions of the good as the first principle rather than his two principles of justice? Rawls would argue that since they already choose the two principles of justice by themselves in the original position, they cannot refuse to accept them. But they may argue that they never accept such principles or anything in the original position at all, why do they need to accept them? Rawls would argue that even though they never actually do so, they must accept them because *if* they had thought of themselves *rationally*, then they would have accepted them.

Now we reach the climax. Rawls’ principles of justice can be justified only if every actual person accepts that his real nature/essence is what described by Rawls’ theory of the person in the original position. To accept that I myself have chosen those principles of justice is to accept that the conception of my real essence as a person has already been

described by Rawls himself. In this sense anyone who believes that he was born not by accident must accept that his comprehensive doctrine is at best the second-order truth; there is something truer than his own conception. Even though they are free to pursue their own conceptions of the good, they must be aware all the time that a theory of the person described in the original position is always truer than their conceptions. Sandel (1982) rightly argues that “we must be prepared to live with the vision contained in the original position, mutual disinterest and all, prepared to live with it in the sense of accepting its description as an accurate reflection of human moral circumstance, consistent with our understanding of ourselves” (p.48).

If Rawls still insists that “*if they had thought of themselves rationally, then they would have accepted them,*” then he cannot deny that people are first required to understand themselves as Rawls described in the original position, otherwise he or she cannot accept that they have chosen principles of justice by themselves; since two principles of justice are chosen by a *deontological* person characterized by Rawls, these principles are just for those who really possess those characteristics. In this sense even though a person holds a strong conception of the good which is incompatible with principles of justice in his private life, he cannot take it too seriously because he must realize that that is not his real nature of the self as his real self must be compatible with principles of justice. Although Rawls would say that they are free to believe that we are not born by accident, he implies that the belief is *less* true than the belief that we are born by accident. But if Rawls insists that they must hold the two principles of justice even though they do not view themselves that way, then his principles of justice would simply be arbitrary, because they must accept principles which they themselves may disagree or have nothing to do with their selves at all. His principles use people as a means.

The only way to get out of this situation is to insist that his principles require people to realize their real selves. To force them to accept two principles of justice is not merely to use them but to make them realize their real selves. In other words, Rawls needs to accept that his theory is

comprehensive/metaphysical, and since his principles of justice (political sphere) are comprehensive, so the 'private sphere' must be determined/constrained by some *particular* comprehensive doctrines, not neutral/political values. It is impossible to separate the 'political' sphere from the 'private' sphere, both political and personal spheres equally require some *particular* moral comprehensive doctrines.

Conclusion

In contrast to Rawls' political liberalism, which holds that it is possible to separate justice and the political sphere from any moral comprehensive doctrines in the sense that principles of justice are not based on any of those doctrines but on political values themselves, I have argued that political liberalism, embraces a *particular* metaphysical/comprehensive theory of the good/person, which is an *unencumbered* self. From a political economy perspective, I have argued that Rawls' principles of justice and the idea of the original position are justified not on the basis of the social contract argument, but on Rawls' theory of the person or the circumstances of justice, which are determined by his own conception of rationality. Without his theory of the person his principles of justice would be impossible because different theories of the person lead to different principles of justice. From a philosophical perspective, I have argued that Rawls' theory of the person in the original position is comprehensive rather than political because its characteristics are given by what Rawls thinks are just in the first place, and that political sphere cannot be separated from private sphere because principles of justice governing political sphere are always *prior to* any other comprehensive doctrines in private sphere.

END NOTES

¹ Rawls sometimes uses the term ‘device of representation’ instead (1993, p. 75).

² To be clear, I think that Rawls’ political liberalism is also metaphysical. But I intentionally use the term ‘comprehensive’ in many places because this should be boarder than merely ‘metaphysical,’ and Rawls would find it harder to deny.

³ Important certain characteristics of the person in the original position are that each individual has the equal moral power to decide the best principle of justice and that they do not know their own social and economic positions and conceptions of good life in the reality until the veil of ignorance is uncovered. Thus, the person in the original position is described as a rational being whose self is *prior* to any end/good (deontological/unencumbered self).

⁴ In fact, Kymlicka (1989) defends Rawls’ political neutrality, but the neutrality and political liberalism share the same aspects so I think we can say that Kymlicka also defends political liberalism.

⁵ She usually uses the term “considerations.”

⁶ This table is revised from Stiglitz & Driffill (2000, p. 381).

⁷ My point here is not to judge whether we are all encumbered or unencumbered selves as Sandel claims. I just want to show that Rawls’ theory of the person cannot avoid this debate, and so his theory must be comprehensive.

⁸ The question is why is physical harm more concerned than spiritual harm? Why is it possible to criticize someone’s beliefs but impossible to attack him physically? This implies that liberals already judge that physical harm is more dangerous/wrong than spiritual/intellectual harm. In other words, since an individual is conceived as an *unencumbered* self who is *prior* to any ends/beliefs, so it is ‘acceptable’ to attack his belief as it is just a *contingent* part of his self, but it is ‘unacceptable’ to attack his body because this is the most fundamental element of his self.

⁹ Kant (1785) actually uses the term the intelligible world or the world of understanding instead of the noumenal world, and the term the world of sense instead of the empirical/phenomenal world. But I think that the noumenal and empirical/phenomenal world are also appropriate here. Please see Sandel (1982).

¹⁰ One may argue that Kant does not defend a rational being on the basis of higher goods/nature of the person. But I would argue that without this assertion, Kant’ theory would be inconsistent. This is because since Kant wants to avoid using people as a means and to respect them as an ends in itself, to say that every rational being must obey the universal moral laws without making it clear that doing so is the way to realize their real selves would be to say that they must do things which they may not agree and has noting to do with their selves at all; it is like imposing the moral laws on people regardless of their consent. To avoid using people as a means and to preserve the moral laws is to propose that complying to the moral laws are things good in itself and good for every rational being. In this way to impose the universal moral law on everyone can avoid using people as a means because even though they may not consent to do so, we can argue that this is the real nature of a rational being which they need to realize, despite

they may not realize it now. In this sense I agree with Taylor (1985b) who argues that “rationality thus imposes obligations on us. Being rational agents, that is possessing a capacity so much higher than nature, puts us under an obligation to live up to this status. We could say that the fundamental principle underlying Kant’s whole ethical theory is something of this form: live up to what you really are – rational agents. Because it is something higher, rationality commands our respect. That is why we experience the moral commands as higher than the demands of nature. We recognize that the moral law is something that commands our respect (*Achtung*)” (p.324).

BIBLIOGRAPHY

Aristotle. *The Nicomachean Ethics*, translated by Ross, D. Oxford: Oxford University Press, 2009.

———. *The Politics*, translated by Ernest Barker. Oxford: Oxford University Press, 1995.

Arrow, K.J. (1951). *Social Choice and Individual Values*, 2nd edition. New York: John Wiley & Sons, 1963.

Bentham, J. (1879). *An Introduction to The Principles of Morals and Legislation*. In Curtis, M. (Ed.). *The Great Political Theories Volume 2: A Comprehensive Selection of the Crucial Ideas in Political Philosophy From the French Revolution to Modern Times*. New York: Harperperennial Modern Classics, 2008, pp. 117-120.

Brown, L. (2009). *Introduction, In Aristotle. The Nicomachean Ethics*, translated by Ross, D. Oxford: Oxford University Press.

Hampton, J. (1980). “Contracts and Choices: Does Rawls Have a Social Contract Theory?” *Journal of Philosophy*, 77(6), 315-338.

Harsanyi, J.C. (1975). “Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls’s Theory.” *The American Political Science Review*, 69(2), 594-606.

- . (1977). “Morality and the Theory of Rational Behaviour.” *Social Research*, 44(4), 623-656.
- Hausman, D.M. and McPherson, M.S. (1996). *Economic Analysis, Moral Philosophy, and Public Policy*, 2nd Edition. Cambridge: Cambridge University Press.
- Kant, I. (1785). *Fundamental Principles of The Metaphysic of Morals*, translated by Abbott, T.K. New York: Prometheus Books, 988.
- . (1788). *Critique of Practical Reason*, translated by Beck, L.W. Indianapolis: Bobbs-Merrill, 1956.
- Kymlicka, W. (1989). “Liberal Individualism and Liberal Neutrality.” *Ethics*, 99(4), 883-905.
- Locke, J. (1690). *Second Treatise of Government*. Cambridge: Hackett, 1980.
- MacIntyre, A. (1981). *After Virtue: A Study in Moral Theory*, Third Edition. Indiana: University of Notre Dame Press, 2007.
- Mulhall, S. & Swift, A. (1996). *Liberals and Communitarians*, Second Edition. Oxford: Blackwell Publishing.
- Rawls, J. (1971). *A Theory of Justice*, original edition. Cambridge: Harvard University Press.
- . (1977). “The Basic Structure as Subject.” *American Philosophical Quarterly*, 14(2), 159-165.
- . (1980). “Kantian Constructivism in Moral Theory.” *Journal of Philosophy*, 77, 515-572.
- . (1985). “Justice as Fairness: Political not Metaphysical.” *Philosophy and Public Affairs*, 14(3), 223-251.
- . (1987). “The Idea of an Overlapping Consensus.” *Oxford Journal of Legal Studies*, 7(1), 1-25.

- . (1988). “The Priority of Right and the Idea of Goods.” *Philosophy and Public Affairs*, 17(4), 251-276.
- . (1989). “The Domain of the Political and Overlapping Consensus.” *New York University Law Review*, 64(2), 233-255.
- . (1993). *Political Liberalism*, Expanded Edition. New York: Columbia University Press.
- . (1999). *The Law of Peoples: With “The Idea of Public Reason Revisited”*. London: Harvard University Press.
- Sandel, M. J. (1982). *Liberalism and the Limits of Justice*, 2nd edition. Cambridge: Cambridge University Press, 1998.
- . (1984). “The Procedural Republic and the Unencumbered Self.” *Political Theory*, 12(1), 81-96.
- . (1994). “Political Liberalism.” *Harvard Law Review*, 107(7), 1765-1794.
- . (2009). *Justice: What’s The Right Thing To do?* London: Penguin.
- Sen, A. (1970). “The Impossibility of a Paretian Liberal.” *Journal of Political Economy*, 72(1), 152-157.
- . (1983). “Liberty and Social Choice.” *Journal of Philosophy*, 80(1), 5-28.
- . (1984). *Collective Choice and Social Welfare*. Amsterdam: Elsevier.
- Stiglitz, J.E. and Driffill, J. (2000). *Economics*. London: W.W.Norton & Company.
- Swift, A. (2014). *Political Philosophy: A Beginners’ Guide For Students and Politicians*, 3rd Edition. Cambridge: Polity Press.
- Taylor, C. (1985a). *Human Agency and Language: Philosophical Papers 1*. Cambridge: Cambridge University Press.
- . (1985b). *Philosophy and the Human Sciences: Philosophical Papers 2*. Cambridge: Cambridge University Press.

———. (1989). *Sources of the Self: The Making of the Modern Identity*.
Cambridge, MA: Harvard University Press

Walzer, M. (1983). *Spheres of Justice: A Defense of Pluralism and Equality*.
New York: Basic Books.