

RESEARCH ARTICLE

A Critique of Contemporary European Liberalism: Why Does Rawls Fail Twice to Justify Legitimacy?

Peining Gao

Peking University

Abstract

This essay aims to criticize contemporary European liberalism in terms of Rawls's hypothetical contractualism. It argues that contractualism is unable to justify legitimacy due to the irreconcilable contradiction between the individuals' plural wills and the sovereign single will. This contradiction lies in Rawls's hypothetical contractualism and turns his view from a Kantian comprehensive conception into an overlapping consensus. In his earlier works Rawls fails to justify legitimacy for the reason of holding a moral dualist view, but later Rawls fails again by ignoring the fragility of an overlapping consensus when the constitutional essentials are at stake. In fact, Rawls's failures are rooted in the idea of contractualism that he shares with Hobbes and Rousseau who also entangle themselves in the same contradiction.

Keywords: Rawls; hypothetical contractualism; legitimacy; priority of right; sovereignty

1 Introduction Since 17th century *contractualism* has been widely believed to be the foundation of *political legitimacy*. But in its critics' view, contractualism overemphasizes individualism and neglects the common value of the state or society. In late 20th century the legacy of contractualism was passed on to John Rawls who endeavored to revive contractualism by developing his hypothetical contractualism to respond to this criticism. Through his work, in particular his dramatic transition, the idea of contractualism reveals its inner tension more clearly than ever before.

For most commentators who are primarily concerned with Rawls's notion of justice, his transition from his earlier works to his later ones often causes a lot of confusions. This transition is sometimes seen as being from Enlightenmentism to Historicism (Rorty 1991, p. 262), or from Neo-Kantianism to Anglo-American Provincialism (Ackerman 1994, p. 376). In this article I shall try to demonstrate that justifying legitimacy is the invariable mission that runs through Rawls's earlier and later thought and argue that, nevertheless, he fails to fulfill his mission. Then I shall point out how Rawls's failure is actually rooted in the irreconcilable contradiction between the individuals' plural wills and the sovereign single will within the idea of contractualism.

In the introduction of *Lectures on the History of Political Philosophy*, Rawls plainly describes his understanding of legitimacy:

"A legitimate regime is such that its political and social institutions are justifiable to all citizens—to each and every one—by addressing their reason, theoretical and practical. This requirement of a justification to each citizen's reason connects with the tradition of the social contract and the idea that a legitimate political order rests on unanimous consent." (Rawls 2008, p. 13)

Rawls relates legitimacy to two ideas: unanimity and contractualism. Unanimity is the essential requirement of legitimacy that demands every citizen, not only the majority¹, to recognize the political regime. Contractualism plays an instrumental role in reaching such unanimous consent. Comparing the two ideas, we can say unanimity is the end and the agreement of contractualism is just the means to reaching it. Rawls never thinks about splitting up the two ideas, whether in his earlier works or in the later ones. This article will follow Rawls's definition of legitimacy as unanimous consent regarding the political regime.² Nevertheless I shall try to argue that it's impossible to reach legitimacy through contractualism. Rawls claims himself to be a successor to the tradition of social contract. In the *Preface of A Theory of Justice* (*TJ* hereinafter), he tells readers he hopes to develop a theory that "is no longer open to the more obvious objections often thought fatal to it" (Rawls 1999a, p. xviii). These fatal objections are not explained in *TJ* but are made clear later in *Political Liberalism* (*PL* hereinafter) when Rawls responds to Hegel's criticism of social contract (See *PL* Lecture VII §10, "Reply to Hegel's Criticism"). According to Rawls, Hegel opposes traditional contractualism primarily because it confuses the state as a general being with the society as a compound of the contingent and private interests. Hegel's point of view may be revealed more explicitly by his criticism of Rousseau's general will:

"The misunderstanding as to the universal will proceeds from this, that the notion of freedom must not be taken in the sense of the arbitrary caprice of an individual, but in the sense of the rational will, of the will in and for itself. The universal will is not to be looked on as compounded of definitively individual wills, so that

these remain absolute.....The universal will must really be the rational will, even if we are not conscious of the fact; the state is therefore not an association which is decreed by the arbitrary will of individuals.” (Hegel 1974 [1840], p. 402)

Rousseau has contended that a legitimate state is supposed to be established based on the general will (universal will), which is agreed to by all the people from the state of nature that is regarded as the Archimedean point in his theory of social contract. For Rousseau both the state of nature and the social contract specify an actual moment in history, albeit we probably shall never be sure of when they happened or will happen. Thus the people participating into the social contract should be historical beings whose interests, needs and inclinations are subject to historical contingencies. A social contract agreement reached in this way will become invalid whenever people’s contingent and private interests change. Therefore a state created by an original contract is in danger of breaking up all the time.³ Rawls totally agrees with Hegel’s rejection of the “historical process doctrines” (Rawls 1996, p. 286) and endeavors to rectify the theory of social contract into a hypothetical one, which he asserts would secure the general form of the state. To avoid the effect of historical contingencies, Rawls no longer presumes the initial situation as one in history but as one in a hypothetical situation, the original position, in which every actual person is represented by a hypothetical party who is free and equal with the other parties. Accordingly, the aim of Rawls’s hypothetical contractualism is not to deal with any actual issues but to set up a series of moral principles as the criterion to judge social institutions. These moral principles are called *the principles of justice*. Thus Rawls expects the principles of justice to have two aims at the outset: generality and agreement. Generality represents the moral constraints imposed upon the particular interests, needs and inclinations that are subject to historical contingencies. On the other hand, since two statuses of subject have been stipulated, the principles of justice are supposed to be agreed to not only by the hypothetical parties but also by the actual persons who, however, are not engaged in the original position. Affirming the principles of justice requires the actual persons to willingly accept the constraints imposed upon their goods. This requirement, which is not included in the traditional theory of social contractualism, is particularly called by Rawls the issue of *stability*. Thus both of these two aims must be reached if Rawls wants to justify legitimacy by means of his hypothetical contractualism. But I shall argue that Rawls fails to reach them in his earlier and later works. His failure in *TJ* results from a Kantian moral dualism, while his failure in *PL* is due to a Hegelian conception of institutions.

2 Rawls and Kant

2.1 Is Priority of Right a Kantian View?

Rawls says, particularly in *TJ*, that the idea of the priority of right is extracted from Kant’s ethics and claims that his hypothetical contractualism represents a Kantian doctrine, but his relation to Kant is interpreted differently by commentators. Michael Sandel considers the priority of right to be of a Kantian “deontological liberalism” (Sander 1982, p. 1). Samuel Freeman denies characterizing the priority of right by deontology but still tags it as a Kantian view (Freeman 1994). Allan Bloom charges that Rawls misuses Kant’s morality as a tool for satisfying men’s desires (Bloom 1975, p.656). Priority of right is one of the key ideas that Rawls insists on throughout his earlier and later works. Whether it is a Kantian view concerns how to understand Rawls’s final goal. In this Section, I focus on explaining its distinction from Kant’s doctrine and why Rawls still calls his hypothetical contractualism a Kantian view.

Although Rawls always states that this idea derives from Kant, in his major books *TJ* and *PL* he doesn’t exposit clearly how he gets the idea. This puzzle is solved in *Lectures on the History of Moral Philosophy*. Thus in the first subsection, my discussion will concentrate on his relevant lectures on Kant. I have to say that, since they drift far from Kant’s text, I shall take them less as a literal interpretation of Kant’s moral philosophy than as a Kantian moral philosophy revised by Rawls. So Rawls’s interpretative deviation will not be treated as misunderstandings but as his distinctions from Kant.⁴

Rawls defines the principles of justice as an idea that corresponds to categorical imperatives (Rawls 1999a, p. 222). However, we will be misled if we regard the categorical imperative precisely according to Kant’s terminology. In fact, Rawls presents a special interpretation of categorical imperative, which is subtly and significantly different from Kant’s own definition. I shall describe Rawls’s interpretation first and then explain why it deviates from Kant’s definition.

Rawls introduces into Kant’s moral philosophy a new idea, the categorical imperative procedure (the CI-procedure), “by which that imperative is applied to us as human beings situated in social world” (Rawls 2000, p. 162). Rawls thinks that Kant never denies the fact that we are beings with needs and desires. So the categorical imperative should not be viewed as a pure form without any contents. Rather, it is supposed to fit

in our happiness. Rawls elaborates four steps of his CI-procedure about how to construct our actual needs and a priori constraints into the categorical imperatives. The standard form of the four steps is as follows.

“(1) I am to do X in circumstances C in order to bring about Y unless Z. (Here X is an action and Y is an end, a state of affairs.) (2) Every one is to do X in circumstances C in order to bring about Y unless Z. (3) Everyone always does X in circumstances C in order to bring about Y, as if by a law of nature (as if such a law was implanted in us by natural instinct). (4) We are to adjoin the as-if law of nature at step (3) to the existing laws of nature (as the se are understood by us) and then think through as best we can what the order of nature would be once the effects of the newly adjoined law of nature have had sufficient time to work themselves out.” (Rawls 2000, pp. 168-169)

Unlike what Kant says, the CI-procedure suggests the categorical imperative is not absolutely free from the hypothetical imperative. Instead, it starts with the hypothetical imperative. Rawls maintains that our natural needs should not be excluded from moral laws like something amoral. Rather, they are both the starting point and the end of the CI-procedure. Satisfying the needs is by itself moral as long as it stays within the constraints imposed by the procedure. Hence one can see that our empirical being situated in the social world is taken for granted by Rawls (see Rawls 2000, p. 151; p. 162; p. 167). Nevertheless, by the CI-procedure Rawls also implies that the natural needs should not be recognized as moral principles by themselves. They are merely materials or contents to be given forms. These forms are viewed by Rawls as external constraints that are completely independent of the natural needs. In other words, “(t)hese constraints are synthetic a priori” (Rawls 2000, p. 250). As moral subjects, we ought to willingly accept these a priori constraints upon our empirical being regardless of what the consequences are. Thus the CI-procedure also affirms us as a priori being, what is called by Rawls the concept of a person as reasonable and rational (Rawls 2000, p. 250).

In Rawls’s eyes, the empirical being and the a priori being constitute Kant’s concept of a complete moral subject. Perhaps we can conversely say that the complete moral subject is split into two subjects. Kant’s ethics also involves two subjects, but Rawls uses the concept of a priori differently from Kant. For Kant the a priori is completely independent of the empirical, while for Rawls their relationship has two aspects. First, the a priori constraints are independent of the empirical needs; second, the constraints are no more than pure forms, thus making no sense if they don’t satisfy the needs. This is not actually Kant’s view, but let me continue to account for Rawls’s interpretation first. According to Rawls, the two subjects are situated in a very tricky situation. On the one hand, they have to depend on each other because neither of them is able to construct the categorical imperative alone. On the other hand, they both retain their independent statuses because neither of them is deduced from the other. However, given their independence Rawls cannot further explain why the a priori subject is at the same time superior to the empirical subject so that the former could impose constraints upon the latter. In the subsequent analysis we will see more clearly that this contradiction runs throughout Rawls’s earlier and later works and becomes the most fatal difficulty in his hypothetical contractualism. Based on the CI-procedure, Rawls tries to take the following step to reach the idea of the priority of right over the goods. He briefly exposit it like this: the right draws the limit; the goods show the points (Rawls 2000, p. 231). Corresponding to the relationship between the two subjects, the right and the goods are complementary, meanwhile, the right constrains the goods. These two features characterize the priority as a lexical one which recognizes the values of the goods but only within the limit the right draws prior to them (Rawls 2000, p. 156; see also Rawls 1996, pp. 173-174). Still the superiority of the right is unanswered. Then Rawls advances what he calls Kant’s moral constructivism. Rawls views the CI-procedure as a moral constructivist procedure by which all the requirements of practical reason, both the pure and the empirical, are represented (Rawls 2000, p. 237). The empirical needs and the a priori moral constrains are both incorporated into the procedure and together work out the categorical imperatives as their outcome. Rawls reaches a conclusion that is obviously contrary to Kant’s: the categorical imperatives are not a priori and the pure practical reason cannot produce its object out of itself, because the CI-procedure has to take the empirical materials into account (Rawls 2000, pp. 250-252). Looking back upon Kant’s text, we will find that not only this constructivist conclusion but also moral constructivism itself is incompatible with Kant’s moral philosophy. Moral constructivism stands for a method that constructs the moral principles out of the empirical needs, however, it unfortunately falls into “the paradox of method” to which Kant strongly objects.

“Now, this is the place to explain the paradox of method in a critique of practical reason: viz., that the concept of good and evil must be determined not prior to the moral law (it would, so it seems, even have to be laid at the basis of this law) but only after it and by means of it (as is indeed being done here).” (Kant 2002 [1788], p. 84)

For Kant the moral law is unconditionally independent of all hypothetical imperatives and should not take any consequences into account. Rawls has always been suspicious of this. Instead, Rawls asserts that the CI-procedure must take account of the consequences and that any moral conceptions putting aside the consequences are simply mad (Rawls 2000, p. 251; see also Rawls 1999a, p. 26). Their divergence on this point leads to all their other disagreements on the categorical imperative. So an examination of the reason for the divergence will be the key to making clear the relationship between Rawls and Kant. In my view, Rawls's suspicion stems from his misunderstanding of Kant's concept of autonomy. Rawls conceives of Kant's autonomy, or the a priori being, as the concept of a person as reasonable and rational. But Rawls neglects the relationship between autonomy and causality, which plays a significant role in Kant's moral philosophy. Kant argues in *Critique of Pure Reason* that everything in the phenomenal world exists in a relation of causality, which is characterized by the fact that every event taking place at a point of time is determined by what happened at a preceding time. Since nothing or nobody at this moment is capable of changing the past, Kant says everything in the phenomenal world is completely controlled by causality (Kant 2002 [1788], pp. 120-121). The hypothetical imperative is an exemplification of the causal control, because it is under the hypothetical condition. This is why Kant asserts that it's impossible to look for the moral law through the goods in the phenomenal world. Along the causal chain, what we can find is still in this chain. If so, it would also be impossible for a person to be a moral subject because as phenomenal beings we are merely animals driven by physical desires and impulses. This leads to an unacceptable consequence that the laws of the jungle would finally replace the moral law as a result of the absence of *duty*. To enable the moral subject, a person must have autonomy, which should be beyond the control of causality and absolutely independent of the influences from the phenomenal world. It is the categorical imperative that represents such autonomy. Unlike Rawls's interpretation, Kant's concept of autonomy, first of all, indicates our relationship with the phenomenal world. It has nothing to do with rationality or irrationality, what Rawls is concerned with, since it is supposed to go beyond causality. Kant says the categorical imperative may never be found through the hypothetical imperatives (Kant 1988 [1785], p. 29). The only way to rescue the moral law and at the same time still keep the phenomenal world is to postulate the autonomy to be a free cause of a causal chain. That means the subsequent events in the causal chain will be the effects determined by my autonomy, while my autonomy is not restrained by the necessity of causality. The postulation of autonomy breaks the physical law of the phenomenal world, but Kant has a good reason for it. The reason is found in *The Critique of Pure Reason*. In the first *Critique*, Kant explicates his epistemological dualism which distinguishes the noumenal world from the phenomenal world. According to his epistemological dualism, the nature is not something alien to us. Rather, it is synthesized by our faculties of sensibility and understanding. So the necessity of causality is unreal to our noumenal self. Kant asserts that one will find that the contradiction between autonomy and nature is merely an illusion as long as he realizes they must be "necessarily united in the same subject" (Kant 2002 [1788], p. 60), i.e. the noumenal self. On the other hand, the unity implies a possible connection between the moral laws and our happiness. Like Rawls's interpretation, Kant indeed admits the fact that we are surely phenomenal beings with natural needs and even have a mandate to meet them. Nevertheless this connection is not a direct one conceivable in the phenomenal world, but an indirect one that may be pursued only in an infinite process assured by another two postulations, the immortality of the soul and the existence of the God (Kant 2002 [1788], p. 146). The two postulations indicate that even if the connection exists in the noumenal world it is still inconceivable in the phenomenal world. So Kant leaves no leeway for Rawls's CI-procedure on this point. It is time to give a brief conclusion to the comparison between Kant and Rawls. Both of them discuss two subjects, the a priori subject and the empirical subject. In accordance with Kant's epistemological dualism, the contradiction between the two subjects is an illusion merely existing in the phenomenal world and the empirical subject will finally show its inauthenticity in the noumenal world. So there is only one subject truly affirmed by Kant. In contrast, putting aside the epistemological dualism, Rawls cannot unite the two subjects only by means of what he calls the moral constructivism, or the CI-procedure. The split of a moral subject brings about the incompatibility between the two features of the priority of right, the moral constraints and the complementarity. Also, it leads Rawls's hypothetical contractualism into an inextricable dilemma: holding on to the constraints from the right at the cost of the complementarity with the goods, or maintaining the complementarity regardless of the constraints. We will subsequently see that Rawls has tried twice to fix the dilemma but the both times failed. Now let us return to the remarks of the commentators. Sandel's view of the priority of right as a Kantian deontology is partly correct in terms of emphasizing the constraints, but it neglects the other feature of the complementarity. Freeman's disapproval of taking the priority of right as an ontological view is correct for taking the complementarity into account, but he doesn't distinguish this idea from a Kantian view. Bloom's criticism of

Rawls's misuse of Kant's morality is sound in the sense that Rawls affirms the independence of the empirical being, but on the other hand improper as Rawls still thinks the moral constraints stem from another world that is independent of the needs. Rawls follows Kant's moral dualist view. Moral constructivism doesn't doubt that the moral constraints, if they make sense, must be given from a world outside the one filled only with natural needs and desires. So Rawls's separation between the hypothetical situation and the actual situation is exactly what he inherits from Kant.

2.2 Rawls's Moral Dualism

In the revised version of *TJ*, in order to distinguish his hypothetical contractualism from Kant's transcendentalism, Rawls adds a new paragraph at the end of §40 "The Kantian Interpretation".

"Kant's view is marked by a number of deep dualisms, in particular, the dualism between the necessary and the contingent, form and content, reason and desire, and noumena and phenomena....His moral conception has a characteristic structure that is more clearly discernible when these dualisms are not taken in the sense he gave them but recast and their moral force reformulated within the scope of an empirical theory. What I have called the Kantian interpretation indicates how this can be done." (Rawls 1999a, pp. 226-227)

In this paragraph Rawls implicitly admits that he borrows Kant's dualist view but separates it from the transcendental aspect. As is well known, Kant's dualism is characterized by the antithesis between the noumenal and the phenomenal worlds. A denial of the transcendental aspect will necessarily lead to a refusal of Kant's dualism. So we should, first of all, figure out what Rawls means by the dualism in an empirical scope. Rawls agrees with Hegel's criticism of the traditional theory of social contractualism that the general form of a social agreement cannot be brought about singly from people's particular interests. But he doesn't want to abandon contractualism as the means to reaching legitimacy, so the original position is put forward for the purpose of revising the traditional social contractualism. In term of being only a revision, Rawls's hypothetical contractualism doesn't intend to remove its very premise that the people's particular interests constitute the foundation of the social agreement. Therefore he postulates the condition of the circumstances of justice only under which the actual persons are possible to enter the original position. The circumstances of justice are featured by the fact that the individuals have "conflicting claims to the division of social advantages" (Rawls 1999a, p. 110). Rawls's circumstances of justice function like a revisionist version of the state of nature.⁵ Nevertheless his revisionist mission requires him to rule out all the contingent factors so as to work out an agreement with the general form. Thus Rawls advances the original position as a supplement to the traditional social contractualism. In the original position, every actual person is expected to be represented by a hypothetical party who is deprived by the veil of ignorance of all the information about people's advantages and disadvantages in the actual society. Only the public knowledge of people's goods and the societal facts is available to the parties. Rawls claims the constraints upon information are reasonable with respect to situating the parties on equal positions, although the parties are rational and disinterested in each other. By virtue of the rational deliberations between the parties, or what Rawls calls reflective equilibrium, the final agreement will be specified as the principles of justice. Since the contingent information has been blocked, the same principles are always chosen upon the public knowledge available (Rawls 1999a, p. 120). And then Rawls's mission will be finished once the actual persons in the circumstances of justice also accept the principles of justice agreed to by the parties. As we shall see later, he encounters great difficulties in his last step. But for now let us keep our eyes on the original position. Rawls says the original position follows Kant's moral constructivism in setting up the distinction between the reasonable and the rational (Rawls 1996, p. 25), which derives from the distinction between the categorical imperative and the hypothetical imperative (Rawls 1996, pp. 48-49). As was mentioned, Rawls doesn't regard the categorical imperative as an a priori idea. The only a priori element in the CI-procedure is the moral constraints. When Rawls distinguishes his hypothetical contractualism from Kant's transcendentalism I believe he means two things. First, the idea of the principles of justice is empirical; second, the veil of ignorance stems from another empirical realm which is distinct from the one of the circumstances of justice. Although the two realms are both empirical, they belong to completely different worlds. By drawing down the veil of ignorance, the inside and the outside worlds of the original position are set apart. The antithesis between the two empirical realms is analogous to the one between Kant's noumenal and phenomenal worlds, in term of splitting the moral constraints from the actual persons. The split doesn't depend on whether the realms are transcendental or not but on whether the moral constraints are independent from the actual world we are living in. When Rawls sets out to pursue a general form of justice by means of blocking all the contingent information, he has in fact affirmed the independence of the moral world. So both Rawls's hypothetical contractualism and Kant's ethics may be called moral dualism in that they recognize the independence of morality. Accordingly, the two realms involve two

corresponding subjects, the actual persons outside the original position (or in the circumstances of justice) and the parties inside the original position. Rawls may decline to conceive of the parties as a sort of subject, because they are merely defined to be the representatives of the actual persons. However, I want to point out that “representative” is a misleading word for describing the parties’ role.

Normally, a representative is supposed to weigh interests in the same way as the actual person authorizes him. In the circumstances of justice, each actual person is presumed to be only concerned with his personal interests. However, once a party is authorized by an actual person, the veil of ignorance will immediately alter the party’s perspective. According to Rawls’s account, the parties have no idea of any actual person’s favorable good and hate to take any risks, so the parties must take all alternative goods into account and treat them equally. In other words, a party has to alter a personal perspective into the public perspective.

Rawls may disagree with attributing any personality that the actual persons don’t have to the parties, because he defines the parties as merely rational beings and disinterested in each other. Nonetheless, the veil of ignorance has changed the parties’ way of weighing interests. An actual person weighs the interests by maximizing his own good no matter how much it would cost other people, while a party should do it by guaranteeing each good an equal opportunity regardless of whether any personal good would be maximized. With the alteration of perspective, we should no longer call the parties the representatives of the actual persons, because they are chasing essentially different interests. So I argue that the alteration forms the parties into a type of subject distinct from the actual persons.

In addition, I also intend to argue that the antithesis between the two subjects is an inescapable consequence of hypothetical contractualism. As we have seen, to avoid the effect from historical contingencies, Rawls’s hypothetical contractualism aims to let each individual recognize the general form of an agreement reached in a hypothetical situation. The hypothetical situation necessarily refers to a corresponding hypothetical subject in contrast to the actual subject. With regard to this antithesis, it is an insignificant issue whether the hypothetical subject has a moral motivation. What really matters is whether the hypothetical subject holds an independent perspective from the actual subject. I maintain that it does, because otherwise the hypothetical agreement would be corrupted by the contingent interests.

In fact, the antithesis is also involved in T. M. Scanlon’s version of hypothetical contractualism which I take as another illustration. Scanlon maintains that since all actual agreements unavoidably involve substantial judgments morality as an ideal is supposed to stem from the hypothetical realm (Scanlon 1998, p. 155). Nonetheless he takes one step back from Rawls’s point of view. His purpose is not to work out any substantial moral principles but to provide an ideal framework of boundary for moral claims. So in Scanlon’s hypothetical contractualism, the veil of ignorance is not adopted because which information is available to all the parties is not his main concern. What he aims to figure out is the reason for reasonable rejections instead of how to reach a reasonable agreement. For him a claim is morally wrong if there is someone who would reject it in accordance with some reasonable principles (Scanlon 2003a).

Like Rawls, Scanlon conceives of a person’s relatively advantaged status as a challenge to morality. He refers to the reasons on the ground of a particular position as personal reasons or “what we owe to each other”, and asserts that the personal reasons cannot provide justifications for reasonable principles. On the contrary, only interpersonal reasons which are produced by “changing places” may offer a basis for reasonable principles (Scanlon 1998, pp. 218-220).

Changing places requires each person to look into the claims from all the other persons’ standpoints rather than to merely hold on to his particular standpoint. Thus we might say the requirement raises an antithesis between the actual subject with personal perspective and the hypothetical subject with interpersonal perspective. As the reasonable rejection stands for an ideal moral notion, morality must keep away from the actual subject. So for Scanlon these two perspectives constitute an antithesis that is similar to Rawls’s.

Many critics like Donald Dworkin (1977, pp. 150-159), Jürgen Habermas (1995, p. 113) and Sandel (1982, pp. 105-110) among others, charge that a hypothetical agreement is invalid because it brings in no interpersonal sanctions. Their charges are not on point due to the neglect of Rawls’s wide and general reflective equilibrium, which has taken the intersubjective views into account.⁶ In fact, Rawls will be misunderstood if one conceives of the original position as a self-conscious situation. The primary characteristic of hypothetical contractualism is the affirmation of the independence of morality.

In order to pursue the general form of morality, moral dualism separates two realms as well as two subjects. However, Rawls’s initial question is how to let the actual persons recognize the general form of the hypothetical agreement. Scanlon’s reasonable rejection is also expected to be finally applied to the imperfect

world (Scanlon 2003a, p. 133). Thus hypothetical contractualism cannot leave these separations as they are. It has to undertake the last task of reuniting the two subjects. In my view, this is the biggest difficulty in hypothetical contractualism and is in fact impossible to resolve.⁷

The difficulty of reunion is called by Rawls the issue of stability, which indicates a sense of justice that “is stronger and more likely to override disruptive inclinations” (Rawls 1999a, p. 398). Since this issue is the last step for justifying legitimacy, it should be consistent with the two aims of legitimacy: generality and agreement. The two aims correspond to the priority of right’s two requirements, the constraints and the complementarity. How to let the reunion meet both the two requirements turns out to be Rawls’s impossible mission. In *TJ*, he tries to build the reunion on the ground of the constraints, whereas the complementarity is dropped. In *PL*, he turns to satisfy the complementarity but, as we will see, leaves the constraints behind.

In accordance with the constraints of justice upon the goods, Rawls argues in the third part of *TJ* that everyone’s personal rational plan is supposed to be inferior to “the larger comprehensive plan that regulates the community as a social union of social unions” (Rawls 1999a, p. 493). The larger comprehensive plan offers the moral constraints but rules out the complementarity. That means the persons in the circumstances of justice are not allowed to make free choices for their goods. Rather, their choices must be circumscribed by the border that justice has drawn in advance.

Sandel is acutely aware of Rawls’s tendency in *TJ* to build a Kantian moral subject (Sandel 1982, p. 49). Rawls commits himself to justifying the unity of the self. However, from the standpoint inside the original position the unity can only be based on a hypothetical subject who is independent of the actual persons’ desires. So even if the desires are recognized, they have to be conceived of merely as subordinate to the hypothetical subject.⁸ In this respect, the larger comprehensive view coincides with Kant’s noumenal self; and the original position, although still empirical, turns out to be equal with Kant’s noumenal world. For the last step, Rawls returns to Kant in the sense of denying the phenomenal self. The Kantian unity of self finally leads to a consequence beyond Rawls’s expectation: we are not allowed to step out of the original position if morality makes sense.

However, this consequence collides with the prerequisite of Rawls’s hypothetical contractualism, the circumstances of justice. Legitimacy requires the general form of justice to be affirmed by each individual under his free will. In case the affirmation can only be maintained by the constraints externally imposed, legitimacy will definitely be unattainable. This assumption of the circumstances of justice, in which everybody holds different comprehensive views, provides the condition for individuals’ freedom. Therefore legitimacy demands the affirmations from both inside and outside the original position.

Rawls then realizes the stability provided by the third part of *TJ* cannot meet the demand. In his later works, he further asserts that any comprehensive doctrine, not only Kant’s philosophy, is unacceptable to the conception of justice. This assertion may seem confusing at first sight but corresponds with Kant’s remarks on political right. In *The Metaphysics of Morals*, Kant contends that the people have an unconditional duty to submit to the legislative head of a state, even in the case when the supreme authority has been unbearably abused, because the head represents the general legislative will.⁹ Kant has denied the possibility of finding the categorical imperative in the phenomenal world. Why does he regard the legislative will of a state as stemming from the pure practical will? Whatever his reason is, Kant’s political view exemplifies that the constraints from moral dualism tend to result in tyranny once it is applied phenomenally. This result is undoubtedly opposite to Rawls’s intention.

Although later Rawls knows that the stability is impossible to be justified inside the original position, he thinks it is only a sectional problem for the framework of his hypothetical contractualism. So in his later works, especially in *PL*, Rawls tries to reunite the two subjects from the standpoint outside the original position and treat the issue of stability more or less independently. What he never realizes is that the difficulty he faces in stability actually arises from the contradiction within the idea of contractualism.

3 The Difficulty in Rawls’s Practical Turn¹⁰

3.1 Overlapping Consensus and Reconciliation

Later Rawls asserts that the justification for his conception of justice has become a practical task (Rawls 1999b, p. 289). “Practical” is a multivocal word. People use it sometimes in contrast with the word theoretical, and sometimes as a synonymy of realistic. So we should first of all clarify what “practical” indicates in Rawls’s usage. In Rawls’s terminology the practical is put in contrast with the metaphysical. He argues that the conception of justice is supposed to avoid the “claims to universal truth, or claims about the essential

nature and identity of persons.” (Rawls 1999b, p. 288) Rawls’s moral dualism distinguishes the perspective of the actual persons in the circumstances of justice from that of the parties in the original position. In *TJ*, in order to reunite the two subjects Rawls contends that the latter perspective stands for a larger comprehensive view of self that constrains the actual persons’ comprehensive views. However, this contention prevents the parties from stepping out of the original position and contradicts the assumption of the circumstances of justice. The circumstances of justice, in which people are considered to hold different and independent comprehensive views, provide the necessary prerequisite for hypothetical contractualism. Justifying legitimacy demands preserving the necessary prerequisite first. So I argue that the practical turn is not caused by anything beyond *TJ*. Rather, Rawls merely switches his standpoint to the circumstances of justice in order to pursue stability. If my interpretation is correct, it would be better to see the practical turn from the view of moral dualism. Although the circumstances of justice, what later Rawls calls *the fact of reasonable pluralism*,¹¹ are defined to be the environment of the actual persons, they still represent a theoretical assumption. It will be a misunderstanding of Rawls’s intention to regard reasonable pluralism as a model of any particular societies existing in history. So the concept of reasonable pluralism is not exposed to criticism from the empirical studies of Americans’ attitudes.¹² Neither does it allude to an inclination for Anglo-American Provincialism as Ackerman says. Furthermore, it is another misunderstanding to view political liberalism as a sort of historicism. Later Rawls doesn’t abandon the antithesis of the two subjects. The parties still remain in political liberalism. Thus Rorty exaggerates the practical turn when he describes later Rawls as a postmodernist philosopher like Martin Heidegger and Jacques Derrida (cf. Fortier 2010). Rawls’s new strategy for stability resorts to the idea of overlapping consensus, in which the conception of justice is viewed as a module in each of the various comprehensive doctrines and separately endorsed from their particular points of view. As a module, the conception of justice plays a very unusual role among the comprehensive doctrines. On the one hand, it is not a comprehensive doctrine itself, but only one part incorporated into them. The comprehensive doctrines are expected to justify the same conception of justice by their own particular reasons as it stems from each of them. By this way, the particular justifications might secure the affirmation of the conception of justice under the independence of the comprehensive doctrines. On the other hand, the contents of justice are formulated independently. That means the doctrines are not allowed to change these contents. The doctrines’ freedom is limited within how to make themselves consistent with justice. In other words, although the affirmation of justice is available through the idea of overlapping consensus, the two subjects still keep their independence. Thus the affirmation of the conception of justice, or stability, has to be maintained in an indirect and conditional way. Given the pluralistic character of the circumstances of justice, the various comprehensive doctrines should not have any agreement, let alone commonly affirm the conception of justice. To justify overlapping consensus, Rawls has to bring in an additional condition that makes the affirmation possible: the enduring free institutions. Rawls claims political liberalism is premised on a closed society whose members “enter it only by birth and leave it only by death” (Rawls 1996, p. 12). He says the people are not assumed to be abstract persons but those who are living in the social world and are going to spend their whole life in it (Rawls 1996, p. 136). According to the four-stage sequence described in *TJ*, after the parties have adopted the principles of justice the actual persons will move to a constitutional convention to work out a just constitution (see Rawls 1999a, pp. 172-173). Although the information about the particular advantages is still unknown in the constitutional convention, the persons are allowed to learn the public knowledge about their society as the rational ground for achieving the most favorable free institutions. Rawls assumes the free institutions used to run compatibly with the other parts of people’s comprehensive (or partial) views over time in history, and then the conception of justice was extracted from the institutions and incorporated into their views as a module.¹³ It consists of the political values that these comprehensive views commonly recognize. These political values and the public knowledge compose the idea of public reason that each individual is supposed to accept during the debate in a public forum. Rawls argues that when basic justice and constitutional essentials are at stake the public reason would provide a basis for reaching a constitutional agreement and thus ensure the legitimacy of a democratic society. Therefore Rawls attains legitimacy by reconciling the two subjects through the institutions. He postulates that the conception of justice was first materialized in the free institutions a long time ago; then by means of a historical process, it is assimilated into the various comprehensive doctrines and obtains their endorsement. The conception of justice involves two independent justifications: reflective equilibrium and overlapping consensus (cf. Scanlon 2003b). The two justifications correspond to the two independent subjects.¹⁴ In *TJ*, Rawls fails to reunite them while keeping their independence. In *PL*, he changes his strategy by building an indirect unity under the historical condition of the free institutions.

What needs to be noticed is that Rawls seems never to doubt the availability of public knowledge in working out the institutions. Whether in the wide reflective equilibrium or in public reason, public knowledge is merely regarded as a constant in pursuing unanimity. Although the disagreement on public knowledge is briefly mentioned by the burdens of judgment, Rawls admits he doesn't intend to seriously discuss it (Rawls 1996, p. 487). But as we will see in the next subsection, the reconciliation of the two subjects may fail as a result of the disagreement on public knowledge.

I think it's necessary to clarify the reconciliation now. Rawls describes it as an idea by which the individuals may realize their free essence through accepting the political and social institutions (Rawls 2000, p. 331). He claims that it is obtained from Hegel's *Philosophy of Right* (see Rawls 2001, p. 3).¹⁵ However, political liberalism is supposed to keep its distance from any kind of metaphysics. So making clear Rawls's relation to Hegel will be helpful in interpreting political liberalism.

In *Philosophy of Right*, Hegel describes the development of the ideas from abstract right, to morality, to ethical life. At the moment of abstract right the institutions enforce the external negative prohibitions on the individuals. Abstract right aims to define personality and has the form of a commandment: *be a person and respect others as persons* (Hegel 1991 [1820], p. 69). But the personality lacks the support from the individuals' subjective will. Another important aspect of abstract right is that only by the right to things is personality able to be realized in institutions (Hegel 1991 [1820], pp. 70-72).

From Hegel's point of view, the right to things stands for a human's supremacy over things. On the one hand, the things are defined to be "something unfree, impersonal, and without rights" (Hegel 1991 [1820], pp. 73-74). On the other hand, the things resist my possession in a sensuous way to manifest their independence. Nevertheless Hegel conceives of its independence merely as something untrue that is sooner or later to be overcome by my will. After overcoming its independence, my will would be infused into the things and turn them into my property. According to Hegel, anything that is external to my will has no end within itself. Thus the idea of property manifests the supremacy of my will over the things (Hegel 1991 [1820], pp.

75-76) and, meanwhile, provides the basis for abstract right.

Hegel specifies morality to be the general right of subjective will. Subjectivity represents the arbitrary aspect of will. The absolute form of subjectivity is conscience that disregards all the external limitations. Hegel asserts that conscience may lead to the good as well as to the evil. The good intends to claim that what it takes as good should be good for everybody. But at the moment of morality, this claim lacks verification from others. So subjectivity disables the good to gain universality. On the contrary, the evil has no intention of going beyond subjectivity at all. It is satisfied with its own particular interests and holds any attempts to pursuing universality in contempt. Particular interests are considered by Hegel as evil because, he thinks, individuals are destined to lead a universal life (Hegel 1991 [1820], p. 327).

Abstract right and morality are finally unified into the idea of ethical life, especially in the state. Ethical life replaces the previous two incomplete moments and becomes the concept of freedom. In the state, abstract right views conscience as its contents; and the subjective good is affirmed objectively by abstract right. Thus Hegel calls the state the actuality of the substantial will which indicates the accomplishment of the two incomplete moments. Accordingly, those antitheses like subjectivity and objectivity, particularity and universality, are both dissolved in the state. Individuals finally realize that only in the institutions, the specific form of the state, they can get rid of their subjectivity and possess their actual essence (Hegel 1991 [1820], p. 196).

In light of the brief description above, we may see that overlapping consensus apparently doesn't fit the idea of reconciliation in Hegel's sense, because there exist no two subjects in Hegel's picture of the state. Abstract right is merely an objective form without any subjective contents, and morality cannot reach universality if it lacks affirmation from the institutions. Neither of them is an independent moment. The genuine subject cannot realize itself until arriving at the stage of the state.

In contrast, Rawls's conception of justice involves both objective moral constraints and subjective support from the parties. And the reasonable comprehensive doctrines are not analogous to Hegel's concept of the good because they don't aim at universality. According to Rawls's account, each comprehensive doctrine is supposed to justify the conception of justice only from its own point of view, having no interest in looking into others' justifications (Rawls 1996, p. 387). On the contrary, Hegel's concept of the good demands the intention of pursuing universality, so when an individual in the moment of morality claims a good he ought to deem his personal reason to be generalized even though its generality cannot be verified for now. However,

each reasonable comprehensive doctrine doesn't care about whether its own reason would be accepted by others. In fact, Rawls's reasonable comprehensive doctrines correspond to Hegel's concept of the evil.

Although Hegel and Rawls both emphasize the institutions' influence on the individuals, the institutions play quite different roles in their doctrines. For Hegel, the institutions cultivate an individual from a subject being into a truly free being. For Rawls, the institutions resemble a bridge that spans the gap between the conception of justice and the comprehensive doctrines but doesn't harm their independence. Hegel's individual accomplishes its actuality in the social world and stay there, while Rawls's individual will still face the task of stepping into the original position.

Rawls expects the idea of overlapping consensus to resolve the issue of stability. Nevertheless stability not only asks for a sense of justice but also requires justice to override the disruptive inclinations. Once historical contingencies upset the balance between justice and the goods, the actual persons must return to the original position to reimpose moral constraints upon the goods. Rawls doesn't, in fact, take this question into account, but it is inescapable if he still wants to justify legitimacy by means of his hypothetical contractualism. We shall discuss this question in the next subsection.

Additionally, it is important to notice that despite the divergence on the role of institutions, Rawls and Hegel share the same basis for institutions: confidence in the availability of knowledge. As was mentioned, whether in the wide reflective equilibrium or in public reason, Rawls always holds on to the agreement on public knowledge. I don't intend to say Rawls maintains that the agreement should be obtained in reality. In fact, he says it is supposed to be pursued in an infinite process (Rawls 1996, p. 385).¹⁶ Rather, I want to point out that the agreement has to presuppose the availability of knowledge.

Similarly, Hegel contends that the institutions are premised on a human's supremacy over things. This contention comes from his famous proposition: *Substance is subject*. With this proposition, Hegel tries to justify how knowledge actualizes itself. Knowledge begins from our immediate consciousness, and then is alienated by objects. At the final stage of self-consciousness, consciousness realizes that objects are no more than its own representation and objects eventually become a moment of self-consciousness. Thus unlike Kant's thing-in-itself, Hegel doesn't think things have any inalienable independence. Even if they are still alien to us temporarily, their independence will finally be dissolved by our knowledge.

Rawls and Hegel both treat knowledge as only a means to actualizing the institutions. However, in my view they are overconfident in knowledge. Once our knowledge and the institutions fail, both Rawls's overlapping consensus and Hegel's idea of the state will be in great danger.

3.2 The Exception

In Hegel's picture, the state represents the free spirit and all the antitheses are supposed to be dissolved in it. No confusions or contradictions should still remain at this final stage. However, his remarks on the state's relation to religion seem very confusing. He says that since religion belongs to conscience the Church will undermine objectivity if it tries to interfere with the state in the name of truth. Hegel insists the state should disregard any interference from the nonpolitical domains, like religion and science, for the sake of defending its own objective rationality (see Hegel 1991 [1820], pp. 290-304).

The confusion arises from why religion and science are still in potential collision with the state. Subjectivity should have been incorporated into ethical life and realized itself as merely an incomplete moment. It should not bear any disruptive inclinations against the state. If it does, that would only imply the state is split into abstract right and morality again. In that case, what religion collides with is not the state, but rather abstract right. However, Hegel doesn't give a further explanation about how the extraordinary situation takes place.

A similar extraordinary situation also shows up in Rawls's political liberalism. Rawls says that public reason aims to provide the basis for legitimacy once the constitutional essentials and basic justice are at stake. It is a very unusual moment. On the one hand, the constitution cannot function well to support the existing political order and so people's agreement on the constitution begins to break down. On the other hand, the state has not entirely fallen into a disorder and people still have the opportunity to assemble in the constitutional convention to reach a new agreement. This extraordinary situation approximates to what Carl Schmitt calls *the exception*.

Schmitt's concept of the exception is helpful to understanding the extraordinary situation here. Schmitt specifies the exception to be a situation that stays on the borderline between routine and disorder. The exception occurs beyond the prediction of rationality, so it cannot be incorporated into the legal system. It

appears as an absolute contingency and “confounds the unity and order of the rationalist scheme”. But the exception is different from anarchy, because the sovereign is still holding the sovereignty in hand. In Schmitt’s view, the exception is brought about by a great and unpredictable crisis, and it is the time for the sovereign to create a new order to deal with the crisis (see Schmitt 1985 [1922], pp. 5-15). However, Schmitt doesn’t think the people may act as a sovereign as a single person does since the people lack the personalistic feature. Although I disagree with Schmitt on this point I shall leave this question aside in this article.

Neither Hegel nor Rawls has taken the issue of the exception seriously. In principle, Hegel denies any kinds of contingency. At the final stage of the state, all contingencies are supposed to be incorporated into the idea of ethical life to be its moments. So Hegel has to either ignore contingencies in reality or hold a conflicting opinion with his whole picture, like his remarks on religion. In contrast, Rawls could not beg the question of the exception since it is the very moment when the constitution obtains legitimacy. Rawls asserts that when the constitutional essentials are at stake the legitimacy will be justified as long as our exercise of political power is endorsed by each individual (Rawls 1996, p. 217). It is necessary to emphasize that the constitutional essentials that are at stake should indicate the old constitutional contents, while the constitutional essentials to be commonly endorsed refer to the new ones. This distinction could be supported by Rawls’s argument in “Reply to Habermas” on citizens’ political autonomy to adjust and revise the constitution as the social environment changes (Rawls 1996, p. 402). Rawls resorts to public reason to rebuild people’s agreement on the new constitutional essentials, whereas without a further examination of the cause for the exception. If public reason is still the same, why are the old constitutional essentials replaced? The old constitutional essentials must have been endorsed in accordance with public reason in the past. Putting the cause for the exception aside, public reason cannot suffice to justify legitimacy alone. So I shall analyze the cause along with Rawls’s route and then point out that overlapping consensus will inevitably break down during the exception. Rawls specifies two parts of public reason: substantive principles of justice and guidelines of inquiry. The first part is included in the conception of justice and the second part is of public knowledge. Let us look into which one may be the direct cause for the failure of the constitutional essentials. According to overlapping consensus, the conception of justice is endorsed by the reasonable comprehensive doctrines in their own points of view. Thus these doctrines’ support for substantive principles of justice is not supposed to be open to the empirical contingencies. The second part, guidelines of inquiry, is described as “presently accepted general beliefs and forms of reasoning found in common sense, and the methods and conclusions of science when these are not controversial” (Rawls 1996, p. 224). In explaining the four-stage sequence, Rawls mentions more details of the public knowledge that the people in the constitutional convention are expected to have.

“But in addition to an understanding of the principles of social theory, they now know the relevant general facts about their society, that is, its natural circumstances and resources, its level of economic advance and political culture, and so on.” (Rawls 1999a, pp. 172-173)

Rawls supposes the persons should take the public knowledge for granted before they reach a constitutional agreement. Nevertheless, the public knowledge is obtained from the empirical world and subject to empirical changes. Not only the divergence on justice but also the disagreement on public knowledge may lead to crisis for human society.

Recall the Great Depression that took place during 1930s and the debate between the New Deal and traditional liberalism. Did people, while being trapped in the crisis, have any agreement on the principles of social theory? Recalling the argument between pro-gun and anti-gun groups, do the two parties share the same relevant general facts about the society?

In fact, whether the people could agree on the public knowledge never depends on their intentions, but on the natural or social environments that are not in their control. As the empirical contingencies stir up the environmental changes, people’s previous knowledge has to be replaced accordingly. A further explanation for the impact of the contingencies on knowledge inevitably involves a broad epistemological discussion that I cannot pursue here. In this article I can only say that the instances in history and reality remind us that the failure of knowledge should be taken into account when the exception is discussed.

If the analysis above is correct, it is the failure of our knowledge that makes the constitutional essentials at stake. In terms of overlapping consensus, since the substantive principles of justice are endorsed separately from public knowledge, the reasonable comprehensive doctrines’ support for the first part of public reason would persist during the exception. However, in terms of reflective equilibrium, public knowledge is still an indispensable element in working out the conception of justice in the original position. The veil of ignorance blocks the information about people’s particular advantages but still leaves the parties the knowledge about

“political affairs and the principles of economic theory” and “the basis of social organization and the laws of human psychology” (Rawls 1999a, p. 119). The parties are presumed to choose the conception of justice based on the relevant knowledge.

Given the replacement of the previous knowledge, the conception of justice should be renewed by another reflective equilibrium. In other words, when the constitutional essentials are at stake it is time for the actual persons to step into the original position.

Habermas criticizes Rawls for his ambiguity on the justification for the conception of justice. Habermas maintains that in contrast to the original position overlapping consensus merely plays a functional role in the justification because it can only provide acceptance, instead of validity, to the conception of justice (Habermas 1995, pp. 121-122). Under overlapping consensus, although justice is endorsed by each reasonable comprehensive doctrine, none of them is concerned with the reasons provided by others. As a consequence, none of these reasons is valid from the perspective of the others. On the contrary, justice is justified in the original position by reflective equilibrium which provides reasons to all the parties. In fact, Habermas’ criticism of overlapping consensus focuses on the contradiction between particular reasons and public reasons. Nonetheless Rawls doesn’t seem to catch Habermas’ point. In his reply to Habermas, Rawls repeats the conformity between the two justifications but begs the question of their contradiction.

The contradiction won’t show up until the exception occurs. In fact, whether overlapping consensus successfully justifies stability depends on whether it allows the actual persons to step into the original position when the constitutional essentials are at stake. Like the priority of right, stability requires not only the complementarity of the goods and the right, but also the moral constraints upon the goods. In political liberalism, Rawls still insists on the priority of right and holds that the comprehensive doctrines should be tailored to meet the restrictions imposed by the conception of justice once they collide with it. Is the priority of right compatible with overlapping consensus when the constitutional essentials are at stake?

Assuming the comprehensive doctrines allow their adherents to step into the original position, the new conception of justice would be surly different from the old one because of the replacement of the old knowledge. Subsequently the previous reasonable comprehensive doctrines would be announced to be unreasonable. All the contents that are in collision with the new conception of justice would then be tailored to fit into it. In fact, the new conception of justice should no longer be treated as a subordinate module. Rather, it becomes the supreme idea of all the comprehensive doctrines. All the other nonpolitical values could survive only if they are recognized by it. As a result, the comprehensive doctrines would lose their independence like in the third part of *TJ*.

Apparently, the comprehensive doctrines should not let their adherents step into the original position. Then it is impossible for the actual persons to genuinely recognize the independence of the parties. That means the conception of justice has to remain as a subordinate module during the exception. Once the crisis puts the constitution in danger, the comprehensive doctrines will never hesitate to sacrifice the conception of justice to preserve themselves. In other words, they would rather let the constitution and the state collapse as the crisis deepens.

Can the constitution and the state be saved on the basis of the old conception of justice? Can the actual persons lay the original position aside and deliberate with the new knowledge to renew the constitutional essentials? If they can, let us ask another question: How is the new knowledge available during the exception? According to Hegel, knowledge stands for human being’s supremacy over things. So knowledge is not completely attainable until the independence of things is overcome. But during the exception, things begin running out of control although people still have opportunity to overcome them (like during the Great Depression, the American economic system was on the brink of collapse but the federal government still had resources to implement the New Deal). This is a special moment when people have only incomplete knowledge. Part of our knowledge is lost because of the exception, but the remaining is still reliable for us to take actions to deal with the crisis. Complete knowledge belongs to the time before crisis when the things are still in control; ignorance indicates the conquest by crisis. Incomplete knowledge stands between complete knowledge and ignorance. It is characterized by our struggle with crisis.¹⁷

Hegel’s state is established based on the availability of complete knowledge at the end of history. However, it is impossible for the people who are involved in the exception to jump out of their present circumstances into the end of history. Rawls also ideally expects the agreement to be reached by public debate in an infinitely long run. But dealing with the crisis calls for a decision to be made in a limited time. That means people have to enter the constitutional convention with divergent opinions. Rawls knows unanimity is

unrealistic to expect and takes majority rule as an imperfect procedural justice to pursue a just outcome (Rawls 1996, pp. 421-422), but the minority opinions are inevitably sacrificed no matter how just the outcome is. As a matter of fact, legitimacy can never be satisfied without the affirmation from the minorities. Furthermore, if the minorities strongly insist on their own opinions against the majority, the constitution and the state would still be put in the danger of collapse.

The failure of overlapping consensus implies that political liberalism returns to the way that Rawls initially attempted to avert. Rawls intended to keep away from historical contingencies by rectifying the traditional theory of social contract into hypothetical contractualism. However, his reliance on the institutions renders an overlapping consensus as fragile as the actual contractualism when facing historical contingencies.

Rawls's two failures in *TJ* and *PL* show that the task to reunite the two subjects is unavoidably undertaken by subjecting one to the other. Accordingly the state would be finally led to either tyranny or anarchy. In my view, Rawls's two failures in justifying legitimacy are due to the idea of contractualism which he has taken for granted at the outset. In Section III I shall analyze why legitimacy cannot be provided by contractualism.

4 Plural Wills Vs. Single Will

In spite of being dissatisfied with traditional contractualism, Rawls never doubts that contractualism is a reliable means to legitimacy. I shall argue that Rawls and some other contractualists, like Thomas Hobbes and Jean-Jacques Rousseau, all misunderstand contractualism. Hegel describes three elements with reference to the concept of contract:

“Since the two contracting parties relate to each other as immediate self-sufficient persons, it follows that (α) the contract is the product of the arbitrary will; (β) the identical will which comes into existence through the contract is only *a will posited by the contracting parties*, hence only a *common will*, not a will which is universal in and for itself; (γ) the object of the contract is an *individual external thing*, for only things of this kind are subject to the purely arbitrary will of the contracting parties to alienate them.” (Hegel 1991 [1820], p. 105)

The first element states that any kind of contractualism ought to start with the condition of particular (arbitrary) wills, which indicate the state of nature or the circumstances of justice. Although hypothetical contractualism endeavors to replace the particular wills, it still has to take them as its starting point. With regard to the latter two elements, contractualism is, however, divided into two types according to whether or not it is aiming at a general object. The first type is represented by Locke-Nozick contractualism. John Locke denies giving the legislative any power that may transgress individuals' natural liberties (see *Two Treatises of Government*, Book II, Chapter XI, §135.). Robert Nozick attempts to justify that the state should be a minimum state whose functions are limited to protection against force, theft, fraud, and so on (Nozick 1974, p. ix). For both Locke and Nozick, the state, if the association can be so called, has no absolute power. Rather, its power derives from the individuals' natural liberties and is supposed to be circumscribed by them. Hence the state has no right to impose external constraints on the individuals' goods, like demanding them to tailor their comprehensive views. The individuals constitute the only subject in Locke-Nozick contractualism. In this case, legitimacy merely depends on the convergence of individuals' particular interests. In Locke's point of view, once the convergence breaks up the people even have a legitimate right to rebel to dissolve the state. Since this type of contractualism doesn't intend to pursue a general object, it is not our major concern. The second type is represented by Hobbes-Rousseau-Rawls contractualism. In contrast to the previous one, the second type attempts to create a sovereign state through a social contract. This intention is obvious in Hobbes and Rousseau's doctrines, while relatively obscure in Rawls's hypothetical contractualism. Although Rawls doesn't intend to justify the origin of the state like the other two contractualists, this topic could still be inferred from his hypothetical contractualism. In *PL*, Rawls says that the concept of legitimacy he tries to justify concerns the proper exercise of the ultimate political power. The ultimate power is not made clear until he asserts, in “Reply to Habermas”, that his doctrine is consistent with popular sovereignty (Rawls 1996, p. 407). Thus we could get a more explicit idea about Rawls's legitimacy. What he is most concerned with never relates to the power of a limited association, like Nozick's minimum state. Rather, he commits himself to a mission that aims to accommodate the exercise of the sovereignty to the individuals' particular wills.

Once the mission is fulfilled, it is also able to serve the purpose of creating a sovereign state because it has presupposed a non-state (or a non-justice) situation as its prerequisite, no matter if the situation is called the state of nature or the circumstances of justice. It will be feasible for the individuals in the circumstances of justice to alienate their entire individual liberties to a sovereign power provided they are convinced that the sovereignty would be exercised in accordance with a conception of justice that is always acceptable to each of them. Is not Hobbes's *Leviathan* created by the sovereign's promise of preserving each individual's physical

life which is considered to be the utmost interest of everybody? In this sense, what Rawls does is no more than replace Hobbes's primary natural laws of preserving physical life with the conception of justice. Of course, one of his distinctions from Hobbes is his defense for democracy. But in terms of absoluteness, the popular sovereignty in Rawls's democratic state is equal with the sovereignty in Hobbes's monarchic state. So Rawls's hypothetical contractualism could also serve to create the absolute power.

Additionally, I don't mean to say that Rawls's hypothetical contractualism may be anticipated to explain the only origin of the sovereign state; instead, it may only justify one of the possible origins if it succeeds. Likewise, Hobbes admits natural force is another way of creating a sovereign state (see *Leviathan*, Chapter 17).

Distinct from the vulnerability of the first type of contractualism, the second one aims to create a sovereign power that is featured by its absoluteness and unlimitedness. A sovereign state is supposed to be in a general form and capable of keeping its own existence away from historical contingencies. However, as the source of the sovereignty, the state of nature or the circumstances of justice are regarded as being exposed to these contingencies. Thus the second contractualism is brought to an intractable question: How to enforce the social contract?

In the civil society, a contract is undoubtedly enforced by the civil law in case any of the contracting parties refuse to or fail to fulfill their duties. As a matter of fact, all civil laws derive from the sovereignty. In other words, it is the sovereignty that enforces all civil contracts. Then can the sovereignty enforce the social contract in the same way? This is the question that most puzzles Hobbes, Rousseau and Rawls.

Hobbes claims that the subjects should abdicate all their individual rights to the sovereign in order to preserve their lives but have the right to defy the sovereign in case their lives are threatened. Rousseau asserts that the general will guarantees everyone will remain as free as before, but it ought to compel whoever behaves disobediently. In *TJ*, Rawls resorts to a larger comprehensive view for imposing the moral constraints upon the goods but then finds it contradicts the circumstances of justice. Later Rawls tries to maintain the independence of the comprehensive views through overlapping consensus but sacrifices the moral constraints.

My answer to this question is negative. Sovereignty cannot be taken as the product of the social contract and, at the same time, as a third party to enforce the social contract. On the one hand, the sovereignty represents the supreme power, so that any attempts at tracing its origin will inevitably subject it to other kinds of power and then conflict with this concept. On the other hand, the idea of contractualism has presupposed the individuals to be the only subject, so any constraints on their particular wills would be inconsistent with this idea. The two notions compose an irreconcilable contradiction.

However, Hobbes, Rousseau and Rawls all endeavor to settle the contradiction by reconciling the two corresponding subjects: the individuals and the sovereign. As the holder of the absolute power, the sovereign is supposed to exercise the sovereignty according to nobody's will other than his own. Any attempts at tracing the sovereign's will are bound to deny his absolute independence. To put it plainly: the sovereign must be treated as an independent person. I should emphasize that a person here is essentially characterized by an independent single will. It doesn't only indicate a single person but also a collective person, like the people in the constitutional convention. Only the collective person is taken into account in this article.

Under the context of contractualism the individuals are supposed to have plural wills that serve for plural goods, while the sovereign has merely a single will in correspondence with the single sovereign power. Finally, all the antitheses facing Hobbes-Rousseau-Rawls contractualism are reduced to the contradiction between the individuals' plural wills and the sovereign single will. If the former override the latter, the state would end up with anarchy; in the contrary case, the state may be led to tyranny. If the individuals hold on to plural wills, any efforts to unite the two subjects are destined to be in vain. This is why Rousseau brings in the concept of *legislator* to change the particular wills.

Rousseau entangles himself in the vacillation between contractualism and the homogeneity of a people. He distinguishes the state of nature from the civil state, and the particular wills from the general will. He says that once the individuals enter the civil state through the social contract their arbitrary particular wills would be soon defined by the general will (*The Social Contract*, Book I, Chapter viii). However, the individuals are supposed to endorse the social contract according to the agreement by their particular wills. How could the effect conversely constrain its cause? Rousseau clearly realizes this paradox and then argues that in order to create a new constitution the effect must become the cause (*The Social Contract*, Book II, Chapter vii). That means the individuals should have accepted the general will even before they agreed to the social contract.

Rousseau hopes the legislator to utilize civil religion to induce the people to accept the general will. Civil religion approximates to Rawls's overlapping consensus, so I don't think it works as far as Rousseau expects.¹⁸ But the legislator is a crucial figure in creating a new constitution and forming a new people. According to Rousseau, in order to avoid the particular wills the legislator should be an alien rather than one of the individuals. The legislator's mission is to replace the individuals' particular wills with the general will. As long as the general will is commonly accepted, the individuals would give up their arbitrariness and make up a homogeneous people. In Rousseau's words, this process produces an outstanding change from a natural person to a moral person, i.e. a citizen.

It is the legislator, not the individuals, who actually ends up the state of nature and creates a sovereign state. Contractualism ought to be abandoned accordingly since it is incapable of getting rid of the influence from plural wills. In fact Rousseau's argument for the legislator alludes to another justification for legitimacy which is based on the homogeneity of a people. A homogeneous people are supposed to pursue a good that is not subject to historical contingencies. Thus the homogeneity of a people needs to be taken into account in rebuilding legitimacy because it is helpful to securing the general form of the state. Nevertheless, under Rousseau's context it is too ambiguous to be understood properly.¹⁹ I shall conclude this article with a specification of two basic aspects of the homogeneity of a people.

First, all the individuals should aim at the same good, or the same comprehensive view in Rawls's words. This aspect will necessarily arouse concern about the common good, which later Rawls and the other liberalist critics have explicitly expressed. I need to point out that this concern is, in fact, caused by moral dualism that imposes independent moral constraints upon the goods. On the contrary, the new comprehensive view I suggest aims to fit morality into the goods. That implies redefinitions of both morality and the goods. Morality should be regarded as the necessary condition for the goods, and the individuals should be cultivated into citizens who desire only the reasonable goods. Thus the unreasonable goods would be rejected by citizens willingly rather than tailored by any external constraints. At the same time, preserving morality would coincide with preserving the reasonable goods. If the new comprehensive view succeeds in making the two redefinitions, morality and the goods are no longer independent of each other, and the only aim of exercising the sovereignty would be to enable each individual to pursue only the reasonable goods.

But moral dualism puts an obstacle to advancing the new comprehensive view. From the perspective of moral dualism, our actual world is merely under the law of the jungle so that any attempts at looking for morality from the goods will finally turn out to be futile. If so, the necessary conditions for the goods are nothing else other than the immoral. Hence the justification for the first aspect of the homogeneity of a people requires a thorough critique of moral dualism, especially Kant's.

Second, all the individuals should affirm the same judgment in the constitutional convention even with incomplete knowledge. The current deliberative democratic theories mainly focus on dealing with the issue of the preferences for goods. Nonetheless, an individual's will depends not only on what goods he wants to chase but also on how he decides to chase them. Even if the people share a common good, it still may be impossible for them to reach a constitutional agreement if the relevant knowledge is lacked. The unavailability of complete knowledge is not seriously taken into account by Rawls and Habermas. Rawls never doubts about the availability of public knowledge even when the constitutional essentials are at stake. Habermas, who blames Rawls for his normative view, argues that the normative content of human right could only be identified in rational discourses, "in which the only thing that counts is the compelling forces of the better argument based on the relevant information" (Habermas 1996 [1992], p. 103). What if no relevant information is compelling enough to convince all?

Rawls and Hegel think the institutions are established based on our knowledge. However, the failure of the institutions during the exception indicates that our knowledge may break down. That means the people have to enter the constitutional convention with their rational disagreement on how to deal with the current crisis. If deliberation is unable to clear the rational disagreement, how is a legitimate decision possible?

Rawls turns to majority rule for help. He views majority rule as an imperfect procedural justice that is expected to lead to the correct result but may also yield an unjust outcome (Rawls 1999a, pp. 74-75). Although Rawls admits the possible injustice, what he is concerned about is still the outcome rather than the procedure itself (Rawls 1996, p. 427). That implies that the minority opinions don't matter at all if a favorable outcome is reachable. The Rawlsian majority rule merely serves as a means to the substantial justice. However, Rawls may not have expected that the democratic procedure could be abandoned once the just outcome or knowledge²⁰ is able to be identified in a non-democratic way. The intellectual elites who claim to

or are believed to monopolize the knowledge are bound to take a dominant advantage to the others. Schmitt even points out that dictatorship may be compatible with democracy in case the dictator claims to educate the people to recognize and express their own will correctly (Schmitt 1988 [1926], p. 28).

Legitimacy should require each individual, whether belonging to the majority or minorities, to listen to others and be open to change his personal opinion in order to reach a final agreement. Thus legitimacy must be premised on a denial of the availability of complete knowledge during the exception. As nobody takes a more advantageous position, everybody's opinion is supposed to be taken into consideration by others.

Thus the second aspect of the homogeneity of a people calls for a critique of the conception of knowledge-dependent institutions that Rawls shares with Hegel. As was mentioned, Hegel's confidence in knowledge comes from his epistemology that is centered on the proposition "Substance is Subject". Likewise, Kant's moral dualism is rooted in his epistemological dualism. A critique of both Kant and Hegel's epistemology should precede the rebuilding of legitimacy.

Notes

1. Rawls's understanding of majority rule as an imperfect procedural justice collides with his definition of legitimacy. I shall explain this point in Section III.
2. It should be noted that the idea of legitimacy in question here is distinct from that in Rawls's article "Reply to Habermas", in which legitimacy is referred to as the imperfect procedural justice of majority rule. Legitimacy requires the affirmation from everybody. In fact a regime that is justifiable only to the majority is still illegitimate to the minorities.
3. Hegel underestimates the complexity of Rousseau's thought. As a matter of fact, Rousseau clearly realizes this difficulty in the general will and attempts to bring in the legislator to fix it. We will return to Rousseau in Section III.
4. I think Rawls is quite aware of his deviation from Kant's text but doesn't want to amend it. He says: "I believe that it is Kant's intention that the CI-procedure represents all such requirements, as far as this can be done. In characterizing his constructivism, I suppose that it does so, but whether it actually does is of course another matter." (Rawls 2000, p. 237).
5. Sandel criticizes the circumstances of justice as an incompatible prerequisite and claims this assumption simply comes from Rawls's common sense about human society (See Sandel 1982, pp. 28-47). In my view, it is indeed incompatible with the justification for the priority of right provided in the third part of TJ, but it doesn't arise from a non-theoretical view. If my interpretation of Rawls is correct, Sandel mistakes the prime question in Rawls's whole scheme. Rawls's biggest concern is not the ultimate foundation for justice but how to let each individual agree with the general form of justice. In case the two questions collide, the former one must give way to the latter. So it is impossible for Rawls to give up the prerequisite if he still holds a contractualist view. I think that is why later Rawls would rather abandon the third part of TJ than sacrifice the recognition of justice from the actual persons.
6. Rawls distinguishes two sorts of reflective equilibrium. The first one involves merely the personal existing judgments, while the second one aims to make a person's judgments conform to all relevant philosophical arguments (Rawls 1999a, p. 43). Rawls contends the fairness of justice is only concerned with the second sort. It asks a citizen to try his best to learn every alternative view that he didn't take into account and then to work out a consistent conception of justice that may be affirmed by all citizens. So Rawls calls the second sort wide and general reflective equilibrium (Rawls 1996, pp. 384-385). Scanlon's hypothetical contractualism adopts this notion as well (see Scanlon 2003, p. 128, fn 3).
7. With regard to this difficulty, Scanlon implies moral education may play a role in helping the actual persons acquire the interpersonal perspective (Scanlon 2003, p. 139). He maintains that moral education is supposed to teach a person in a particular position changing places to learn what justifications other people accept. But this process will result in a paradox that once the interpersonal perspective is adopted it will become the actual perspective. According to Scanlon's postulation, the actual perspective is defined by personal perspective that only relates to a person's particular position. Thus the consequence of moral education would contradict this postulation. I have no disagreement with moral education and changing places. What I want to emphasize is the irreconcilable contradiction between the personal and the interpersonal perspectives. This point will be explained in more detail in Section III.
8. "Thus in drawing up our plan of life we do not start *de novo*; we are not required to choose from countless possibilities without given structure or fixed contours. So while there is no algorithm for settling upon our good, no first-person procedure of choice, the priority of right and justice securely constrains these deliberations so that they become more manageable." (Rawls 1999a, p. 494)
9. "A law that is so holy (inviolable) that it is already a crime even to call it in doubt in a practical way, and so to suspend its effect for a moment, is thought as if it must have arisen not from men but from some highest, flawless lawgiver; and that is what the saying 'All authority is from God' means. This saying is not an assertion about the historical basis of the civil constitution; it instead sets forth an Idea as a practical principle of reason: the principle that the presently existing legislative authority ought to be obeyed, whatever its origin." (Kant 1991 [1800], p. 130)
10. I borrowed the phrase "practical turn" from Patrick Neal (1994).
11. With respect to the connection between the circumstances of justice and reasonable pluralism, the following text is illuminating: "Since we want the idea of such a society to be suitably realistic, we assume it to exist under circumstances of justice. These circumstances are of two kinds: first, there are the objective circumstances of moderate scarcity, and second, the subjective circumstances of justice. These latter circumstances are, in general, the fact of pluralism as such; although in a well-ordered society of justice as fairness, they include the fact of reasonable pluralism." (Rawls 1996, p. 66).
12. George Klosko opposes the concept of reasonable pluralism by showing the empirical studies which manifests that tolerance is not widely accepted among American people (see Klosko 1993).
13. See the process from constitutional consensus to an overlapping consensus (PL, Lecture IV, § §6-7).
14. The independent status of comprehensive doctrines is favored by a number of liberalists because it avoids the idea of common good. For these liberalists, a common good would be bound to bring about the suppression of the other goods, thus public reason preserves pluralism that is considered to be the essential value of liberalism. Among others, see Galston 1995; Larmore 1999; Gutmann 1995.
15. Joshua Cohen contends that political liberalism does have some important similarities with Hegel's philosophy but should not be viewed as based on the latter (see Cohen 1994).

16. On this point I was indebted to Scanlon's article "Rawls on Justification" (Scanlon 2003b).
17. A full explanation for incomplete knowledge in fact asks for a new understanding of the concept of knowledge that is distinguished from the epistemological tradition since Descartes. But a discussion on it is far beyond the scope of this article.
18. Rousseau asserts tolerance is the only limit that is supposed to be imposed on the various religions (The Social Contract, Book IV, Chapter viii). However, if religious tolerance, along with the general will, has to be supported by these religions, then it will be impossible to be considered as a limit to them.
19. As one of the misunderstandings, Schmitt's criticism of democracy focuses on that the Rousseauian people lack the decisionistic and personalistic elements (Schmitt, 1985 [1922]). But I cannot pursue the defense of these elements in this article.
20. Rawls refuses to equal justice with truth in the original position, but in the constitutional convention justice corresponds with truth in terms of correctness. The original position represents a pure procedural justice that doesn't take any outcome for granted (see Rawls 1999a, p. 104, p. 118). Nevertheless, at the stage of constitutional convention justice has been given specific and substantial contents, so people are actually pursuing a definite outcome in working out the constitutional essentials. Rawls conceives of the democratic procedure in the constitutional convention merely as a means that "is more likely to arrive at the correct conclusion" (Rawls, 1999a, p. 315). Apparently, those people who hold more knowledge are closer to the just outcome than others. Rawls's justice is not so far from truth as he believes.

References

- Ackerman, Bruce. 1994. Political Liberalisms. *The Journal of Philosophy* 91: 364-386.
- Bloom, Allan. 1975. Justice: John Rawls Vs. the Tradition of Political Philosophy. *The American Political Science Review* 69: 648-662.
- Cohen, Joshua. 1994. A More Democratic Liberalism. *Michigan Law Review* 92: 1503-1546.
- Fortier, Jeremy. 2010. Can Liberalism Lose the Enlightenment? *The Journal of Politics* 72: 1003-1013.
- Freeman, Samuel. 1994. Utilitarianism, Deontology and the Priority of Right. *Philosophy & Public Affairs* 23: 313-349.
- Galston, William. 1995. Two Conceptions of Liberalism. *Ethics* 105: 516-534.
- Gutmann, Amy. 1995. Civic Education and Social Diversity. *Ethics* 105: 557-579.
- Habermas, Jürgen. 1995. Reconciliation through the Public Use of Reason: Remarks on John Rawls's Political Liberalism. *The Journal of Philosophy* 92: 109-131.
- Habermas, Jürgen. 1996 [1992]. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy* (trans: William Behg). Cambridge: MIT Press.
- Hegel, G.W.F. 1974 [1840]. *Hegel's Lectures on the History of Philosophy, Volumn III* (trans: E. S. Hardane and Frances H. Simson). New York: The Humanities Press.
- Hegel G. W. F. 1991 [1820]. *Elements of the Philosophy of Right*, ed. Allen Wood (trans: H. B. Nisbet). Cambridge: Cambridge University Press.
- Kant, Immanuel. 1988 [1785]. *Ground of the Metaphysics of Morals* (trans: Mary Gregor). Cambridge: Cambridge University Press.
- Kant, Immanuel. 1991 [1800]. *The Metaphysics of Morals*. (trans: Mary Gregor). New York: Cambridge University Press.
- Kant, Immanuel. 2002 [1788]. *Critique of Practical Reason* (trans: Werner S. Pluhar). Indianapolis: Hackett Publishing Company, Inc.
- Klosko, George. 1993. Rawls's "Political" Philosophy and American Democracy. *American Political Science Review* 87: 348-359.
- Larmore, Charles. 1999. The Moral Basis of Political Liberalism. *The Journal of Philosophy* 96: 599-625.
- Neal, Patrick. 1994. Does He Mean What He Says? (Mis)Understanding Rawls's Practical Turn. *Polity* 27: 77-111.
- Nozick, Robert. 1974. *Anarchy, State, and Utopia*. Oxford: Blackwell Publishers Ltd.
- Rawls, John. 1996. *Political Liberalism (Expanded Edition)*. New York: Columbia University Press.
- Rawls, John. 2008. *Lectures on the History of Political Philosophy*, ed. Samuel Freeman. Cambridge: The Belknap Press of Harvard University Press.
- Rawls, John. 1999a. *A Theory of Justice (Revised Edition)*. Cambridge: The Belknap Press of Harvard University Press.
- Rawls, John. 1999b. Justice as Fairness: Political not Metaphysical. In *Collected Papers*, ed. Samuel Freeman, 388-414. Cambridge: Harvard University Press.
- Rawls, John. 2000. *Lectures on the History of Moral Philosophy*, ed. Barbara Herman. Cambridge: Harvard University Press.
- Rawls, John. 2001. *Justice as Fairness: a Restatement*, ed. Erin Kelly. Cambridge: The Belknap Press of Harvard University Press.
- Ronald, Dworkin. 1977. *Taking Rights Seriously*. Cambridge: Harvard University Press.
- Rorty, Richard. 1991. The Priority of Democracy to Philosophy. In *Objectivity, Relativism, and Truth: Philosophical Papers*, 257-282. London: Cambridge University Press.
- Sandel, Michael. 1982. *Liberalism and the Limits of Justice*. New York: Cambridge University Press.
- Schmitt, Carl. 1985 [1922]. *Political Theology: Four Chapters on the Concept of Sovereignty* (trans: George Schwab). Cambridge: MIT Press.
- Schmitt, Carl. 1988 [1926]. *The Crisis of Parliamentary Democracy* (trans. Ellen Kennedy). Cambridge: MIT Press.
- Scanlon, T. M. 1998. *What We Owe to Each Other*. Cambridge: The Belknap Press of Harvard University Press.
- Scanlon, T. M. 2003a. Contractualism and Utilitarianism. In *The Difficulty of Tolerance: Articles in Political Philosophy*, 124-150. Cambridge: Cambridge University Press.
- Scanlon, T. M. 2003b. Rawls on Justification. In *The Cambridge Companion to Rawls*, ed. Samuel Freeman, 139-167. New York: Cambridge University Press.