The Fundamental Interests of Citizens: A Response to Chung[1]

Author: Justin P. Holt[2]

Abstract: Hun Chung's recent article "Rawls's Self-Defeat: A Formal Analysis" argues that the selection of results equivalent to justice as fairness can be derived by utilitarianism. Chung argues that these results can be achieved through the use of Rawls's constructed utility function from his work *Justice as Fairness*. Although Chung's article is finely argued and presented in great detail, this paper will show that Chung made three mistakes in the fundamentals of his argument. First, Chung mistakes Rawls's constructed utility function as actually utilitarian. Second, Chung confuses the motivations and the assumptions of those held by agents in the original position and those held by utilitarian agents. Last, Chung's previous mistakes result in the incorrect conclusion that an equal distribution of resources will be selected by representative agents instead of the unequal distribution prescribed by the difference principle.

1. Introduction

Hun Chung's article "Rawls's Self-Defeat: A Formal Analysis" provides a well-argued and elaborate reconsideration of how principles of justice are selected within the parameters of Rawls's assumptions on choice. The stated purpose of Chung's article is to show that according to Rawls's own tools the preferred result of the deciding agents is the selection of utilitarian outcomes (Chung, 2018, 4). This is done, rather elegantly, by the use of Rawls's constructed utility function from *Justice as Fairness* (Rawls, 2001, 107-108). Nonetheless, I find Chung's understanding of the use of Rawls's utility function to be incorrect. In particular, this is the case because Rawls's construction of a utility function is not based on utilitarian assumptions. If Chung misunderstands what Rawls's utility function is supposed to show, then his conclusion

---

[2] Institutional affiliation: Wright College in Chicago, IL. Professional email: justinpholt@gmail.com.

that utilitarianism provides better outcomes for the disadvantage is no longer definitively clear. At best Chung perhaps has demonstrated a conclusion (which G.A. Cohen argued for in *Rescuing Justice and Equality*) that equal distributions are of greater benefit to the disadvantaged than the unequal distributions that the difference principle endorses.

In order to show that Chung has misunderstood Rawls's constructed utility function, I will first review Rawls's discussion and use of the function. Next, I will argue that given how Rawls constructed the utility function, Chung has made a mistake in utilizing it as a utilitarian selection device to derive results that are similar to justice as fairness. Finally, Chung is correct that an equal distribution of resources is of greater benefit for the disadvantaged than the unequal distribution that the difference principle encourages. However, if Chung is mistaken about the motivations and assumptions of the agents in his model, then he is incorrect that an equal selection will be chosen by these agents.

2. Rawls's Constructed Utility Function

The best place to begin is to review Rawls's comments on his constructed utility function in *Justice as Fairness*. Rawls introduced his utility function, on the one hand, "to clarify the relation between justice as fairness as a political conception and utilitarianism as a comprehensive view" (Rawls 2001, 107 note 30). On the other hand, this consideration of utilitarianism becomes an issue because the third condition for the use of the maximin rule in selecting the principles of justice holds that a guaranteeable level, which the conception of justice as fairness is supposed to provide, is desirable because "the worst outcomes of all the other alternatives are significantly below the guaranteeable level" (Rawls, 2001, 98). Rawls's use

of his constructed utility function is to respond to the objection that "the argument from the third condition is utilitarian after all" (Rawls, 2001, 107). Rawls argued that justice as fairness is not utilitarian, and his response is to show that advocates of utilitarianism as a comprehensive doctrine "can join in an overlapping consensus on that conception [of justice as fairness]" (Rawls, 2001, 109).

Rawls wished to show that utilitarians can reasonably agree to the basic structure. This is the same as showing a Christian that it is possible within a justice as fairness society for them to practice Christianity in fundamental ways: constructing churches and schools, worker protection for their Sabbath day, allowance for public Christian festivals, and substantial preservation for their believes in many areas, including birth, death, marriage, and education (this list is not meant to be exhaustive). Any politically liberal society should be open to the inclusion of comprehensive doctrines that will not violate the principles of justice, this includes utilitarianism. A political liberal society should be able to show an adherent of a tolerant comprehensive doctrine that they too can live here and hold their ideas of ultimate value among others who embrace different comprehensive doctrines that are also tolerant.

All in all, Rawls showed that utilitarians can find that their ultimate system of value can be practiced within a society where justice as fairness is the political conception that organizes the basic structure. His presentation is not to show that justice as fairness is utilitarian. In seeking these ends Rawls conceptualized the constructed utility function as such:

I hold that it is not [justice as fairness is not utilitarian]. This is because the parties use a utility function (I am willing to call it that) so constructed as to reflect the ideal normative

conceptions used to organize justice as fairness, that is, the ideas of society as a fair

system of cooperation and of citizens as free and equal, and characterized by the two

moral powers, and so on. This constructed utility function is based on the needs and

requirements of citizens – their fundamental interests – conceived as such persons; it is

not based on people's actual preferences and interests. (Rawls, 2001, 107)

The utility function that Rawls constructed uses particular assumptions to build its schedule.

First, the function reflects the norms which characterize justice as fairness. Second, the agents

who select this schedule are citizens, which means agents under the veil of ignorance in the

original position, not people with "actual preferences and interests." These assumptions were

fundamentally important for Rawls's inclusion of utilitarianism within the overlapping consensus

of a political liberal society that takes justice as fairness as its political conception. Rawls wanted

to show that utilitarians can agree to the principles that are decided within the original position

under the veil of ignorance as utilitarians. Rawls did not show that justice as fairness can be

derived from utilitarian premises.

What utilitarians are agreeing to is that the demonstration of the principles of justice as

fairness can include utilitarian notions of value, not that justice as fairness is based on utilitarian

notions of value. Utilitarians agree *within* the suitable constraints of the two assumptions above.

The principles of justice are decided first on their own since they are part of a free-standing

political conception that is not reducible to any particular comprehensive doctrine. Utilitarian

value reasoning can overlap with the minimal standards of the basic structure.[3] Second, it is

shown that the kind of utility function that would overlap will be one that uses the "needs and

requirements of citizens" and not "people's actual preferences and interests." This means that the

utility function that Rawls constructed is based on choices of agents under the strictures of the

veil of ignorance within the original position and not based on the choices of agents who accept

utilitarian standards of value.

The utility function that Rawls used is not based on utilitarian reasoning for its

construction. Rather, it is an expression of the principles of justice in a utilitarian formulation.

Rawls described this feature as such:

> The reply to the utilitarian, then, is that justice as fairness does not deny that the idea of a
>
> utility function can be used to formulate justice as fairness. Indeed, I suppose any
>
> conception of justice can be expressed as maximizing a suitably doctored utility function.
>
> Surely the debate between utilitarian and nonutilitarian views is not about that trivial
>
> formal question! Such a utility function is but a mathematical representation that encodes
>
> certain basic features of our normative assumptions. That there is a representation of this
>
> kind says nothing about the content of justice as fairness; nor is it the substantive idea of
>
> utility in the tradition of utilitarianism. (Rawls, 2001, 107)

---

[3] Rawls stated also that "… we should be cheered if utilitarians can find, from within their own point of view, a way to endorse the ideas and principles of justice as fairness" (2001, 107-109). A justice as fairness society can show that your doctrine can overlap with the free-standing political conception, but this is not forcing members who adhere to a comprehensive doctrine to accept membership. They must consider if their conscience allows them to live in a society where their comprehensive doctrine is not the political conception of the basic structure. If not, the best we can hope for is a *modus vivendi*.

Rawls's constructed utility function is merely an expression of justice as fairness, with the usual terminology of utilitarianism. Nonetheless, these principles have not been determined with utilitarian tools and terminology. Also, there is no attempt to translate the principles of justice as fairness into utility. Rather, this is a presentation of justice as fairness for comparison to utilitarianism, not reduction of one to the other.

There are many differences with utilitarianism here. In particular, a utilitarian metric of evaluation – preferences, interests, or pleasure – is displaced by the principles of justice. Preferences are not maximized, rather it is argued that as justice as fairness, conceived as a guaranteeable level, can be "doctored" into a utilitarian function. Instead of a utilitarian construction of justice as fairness, this appears to be a demonstration of public reason from the standpoint of an advocate of justice as fairness (Rawls, 1996, 23 and 45). It is not a demonstration from the standpoint of a utilitarian utilizing public reason. If a utilitarian were discussing how their utilitarian system of value could exist as a tolerable doctrine, among other tolerable doctrines, then there would be a demonstration of utilitarian values via public reason.

3. Chung's Reading

Given Rawls's presentation of his constrained utility function, I find that Chung has missed a key distinction when he argues that utilitarianism can be shown to produce results similar to justice as fairness. Chung states:

> … I would like to emphasize that the only reason why we are considering a liberal
>
> democratic society that meets the principle of equal basic liberties (i.e. Rawls's first

principle) is to meet the preconditions that allow us to apply Rawls's difference principle. Such an assumption is not meant to restrict the distributional consequences of utilitarianism. All the distributional results of utilitarianism that we will soon derive will remain intact even if we dropped this assumption. So, what we are really comparing is the distributional consequences of utilitarianism-full-stop (not utilitarianism restricted by Rawls's first principle) and Rawls's difference principle. (Chung, 2018, 16)

In short, Chung's model is not simply a consideration of utilitarianism replacing the difference principle. Rather, Chung's model is considering utilitarianism compared to the entirety of justice as fairness. Still, Chung assumes that the first principle of justice shapes an agent's choices. This means that the assumptions and motivations of those who are choosing principles of justice in his model are not really utilitarian agents. As discussed above, these are agents acting as citizens using public reason to appeal to utilitarians. Also, it seems that Chung does not make good on the claim that the results will hold if the assumption is dropped concerning the first principle. If this was the case, then this sounds as if the basic assumption of utilitarian choice is not the maximization of summed preferences. Rather, Chung is assuming that utilitarian agents select the maximization of utilities after the first principle has been settled. This looks like a restricted utilitarianism or a mixed conception, but Chung does note in his paper that his theory is not of this type (Chung, 2018, 20). I find that this matter is confused in his presentation, since there is little clarity in defining the difference between the motivational assumptions of utilitarian agents and Rawlsian agents.

To begin, it is good to note that the difference between "the needs and requirements of citizens" and "people's actual preferences and interests" has been attended to in Chung's reading:

> … the difference principle is completely blind to the issue of what different bundles of primary social goods *can actually do for different people*; … In contrast, utilitarianism *does* care about what different bundles of primary social goods do for different people; it cares about how these different bundles of resources translate to people's welfare and attempts to maximize their total sum.[4] (Chung, 2018, 18).

This is true; but, if the utility function that Rawls constructed is not based on people's actual preferences, then it seems that the utilitarianism that Chung is analyzing is not based on actual people's welfare states. Rather, Chung is analyzing the fundamental interests of citizens, which are agents constrained in their choices by the veil of ignorance within the original position. This is a mistake in specifying what are the motivations and assumptions of those who are selecting

---

[4] Chung notes Sen's critique of social primary goods as resource fetishism (Chung, 2018, 18). Rawls responded to these claims in section 51 of *Justice as Fairness*. In this section, Rawls considered the flexibility of social primary goods. Rawls noted that in the legislative stage of his theory there is space for flexibility in the distribution of primary goods. Interestingly, this flexible distribution still utilized the notion of a citizen: "Important here is the use of the conception of the citizen as a cooperating member of society over a complete life, which enables us to ignore differences in capabilities and endowments above the minimum. That conception directs us to restore, or in an appropriate way to make good, our capabilities when by illness and accident we fall below the minimum and are unable to play our part in society" (Rawls, 2001, 175). Thus, Rawls found that the representative agent of the original position can be used as a flexible guide in the distribution of primary goods. This consideration by Rawls calls into question Chung's distributive comparisons. In particular, the difference in distributions between the most advantaged and the least advantaged may be quite different when spending like health care is taken into account. Citizens may receive equal access to health care and equal treatment, but they would not receive equal care since this would depend on the care they needed. For example, not all people will have their broken bones reset, since not all people will break their bones. Also, those who do not break their bones would not be compensated with resources that were used for those who did break their bones. Unfortunately, there is not room in this paper to consider this section in *Justice as Fairness* further.

principles of justice. Fundamentally, Chung is confusing the selection of the principles of justice

from a utilitarian starting point and Rawls's principles of justice formulated as a utility function.

As a case in point, Chung notes:

> Remember that the reference point of each group's utility function is supposed to denote
>
> the amount of primary social goods (in this case, wealth) each member of the group needs
>
> to fully enjoy *the equal worth* of the full set of basic rights and liberties that is formally
>
> guaranteed by our model society meeting Rawls's first principle of justice. (Chung 2018,
>
> 14)

If these representative agents are seeking primary social goods for the two principles to be in

effect, then their choices are being shaped by the needs and requirements of citizens. This is

inconsistent with "what different bundles of primary social goods *can actually do for different*

people." There are no actual people in the original position; if there were, then there would be

the problem of egoism to confront. The veil of ignorance sifts out these positions (Rawls, 1999,

117).

As was shown above, Rawls provided a constructed utility function not to show that

justice as fairness is utilitarian. Instead, this is done to show that "a utility function can be used to

formulate justice as fairness." This is different than saying that utilitarian assumptions derive the

principles of justice as fairness. Utilitarianism seeks the maximization of actual people's

preferences, not the demonstration of what an agent would select under the conditions of the veil

of ignorance in the original position. These are two different rational subjects, and accordingly

they would generate different principles of justice. Chung intends to use the rational subject

whose choices reflect the strictures of the veil of ignorance in his demonstration when he follows

Rawls's presentation. This means that Chung's interpretation is not really utilitarian, but rather is

justice as fairness in utilitarian clothing. The principles of justice are formulated under the veil in

the original position first and then they are presented with a utilitarian appearance.

If Rawlsian principles are supposed to be selected via a utilitarian agent, then this is not

shown by Chung. There is an unclear mixing of the choices of agents under the veil in the

original position and agents who have utilitarian choice schedules. If agents were really choosing

according to the demands of utilitarian choice schedules, then they would not choose the political

conception of justice as fairness. This is the case because justice as fairness is not a conception of

justice that attempts to maximize utility summed over an aggregate group. Thus, Chung confuses

utilitarian agents with Rawlsian agents in his demonstration.

4. Equal Distribution

Even given these shortcomings Chung's work is rather impressive and attempts to support an

egalitarian reading of the difference principle. It is demonstrated that an egalitarian distribution is

more of a benefit to the disadvantaged than an unequal distribution. As Chung points out,

Rawls's argument is to avoid an egalitarian distribution (Chung, 2018, 13). Nonetheless, the

utilitarian distribution that Chung demonstrates is egalitarian in its outcomes:

> The way utilitarianism distributes the available resources accords very well with our
>
> basic moral intuitions. It prescribes, first, to secure the equal worth of everybody's basic

rights and liberties by giving everybody the amount of resources needed to satisfy his/her

specific reference point. Afterwards, utilitarianism prescribes to divide and distribute the

remaining social wealth—i.e. what is left after everybody secures the equal worth of

his/her basic rights and liberties—equally to everybody. (Chung, 2018, 19)

In order to reach this conclusion, there has to be an alteration of the incentives that are

encapsulated in the difference principle.[5] Rawls assumed that incentives were necessary in order

to benefit the disadvantaged. As Cohen (2008) showed this use of incentives does not really

result in the greatest benefit to the disadvantaged. Cohen's analysis was to consider if these

incentives are really necessary for increased output. Cohen identified that there are morally

legitimate reasons for incentives, such as, compensation for arduous or undesirable work and for

the costs of education and training. Nonetheless, incentives may serve as an excuse for the more

advantaged members of society to extract a premium in remuneration that does not reflect

necessary compensation.

Chung's argued result of an equal distribution does not consider the real depths of why

we assume that incentives are necessary in order to obtain greater amounts of time spent at work,

greater output, or highly trained and educated workers. As Cohen focused on, why do people

who work in comfortable environments and perform interesting work need additional

compensation? Rawls embraced incentives seemingly because it is generally assumed in the

social sciences that incentives are necessary to encourage people to perform labor, given the

---

[5] Also, there appears to be prior assumptions or moral intuitions that are shaping the selection of equal basic liberties here. These agents are not just maximizing summed preferences. Rather, they are, either, calling equal basic liberties maximized summed preferences. Or, these agents have moral intuitions that consider certain outcomes as having greater weight in the satisfaction of preferences than other outcomes.

consideration that labor is a pain or a bad. This is far from the truth for those who are engaged in highly trained professions.

These considerations of incentives are unexamined in Chung's paper, and this is a problem with the confusion between the choices of Rawlsian agents and utilitarian agents identified in the preceding section. It is assumed that Rawlsian agents require compensation to provide the greatest benefit for the disadvantaged. Rawls thought this was a demonstration of reciprocity (Rawls, 2001, 49). Whereas, utilitarian agents, following Rawls, are more willing to perform acts of benevolence (Rawls, 1999, §30). The problem is identifying if Chung's agents are utilitarian or Rawlsian in their motivations. If these agents, as I argued above, are motivated in the ways the Rawls argued, then they are not motivated in ways that would result in an egalitarian distribution. Rawls noted this in *A Theory of Justice*:

> From the standpoint of justice as fairness there is no reason why the persons in the original position would agree to the approvals of an impartial sympathetic spectator as the standard of justice. This agreement has all the drawbacks of the classical principle of utility to which it is equivalent. If, however, the parties are conceived as perfect altruists, that is, as persons whose desires conform to the approvals of such a spectator, then the classical principle would, of course, be adopted. (Rawls, 1999, 164)

This means that Chung's egalitarianism, in the guise of utilitarianism, may not be a definitive result in his model if the motivations of the agents in his model are not clear.

5. Final Comments

Chung has written a detailed and engaging essay that should make us reconsider Rawls's criticism of utilitarianism. This is a timely piece since utilitarianism and consequentialism, in general, have been called into question as suitable normative theories over the past several decades. As I showed above, I do think there are substantial problems with the foundations of Chung's analysis that call into question his conclusions. In particular, it appears doubtful to me that one can consider his presentation to be actually utilitarian given that Rawls's constructed utility function is an expression of justice as fairness. If the limits on the range of preferences that can be selected by agents in a model are tightened, then the further away one moves from being able to hold that this theory is truly a maximization of summed preferences (Nussbaum, 2000, 122-128). Chung wishes to argue that his presentation is not a form of restricted utilitarianism or a mixed conception (Chung 2018, 20). This, in a way, is true. Simply, Chung's presentation is really justice as fairness, with associated mistakes involving the motivations and assumptions of agents in the original position.

References

Chung, H. (2018, December 7). Rawls's Self-Defeat: A Formal Analysis. *Erkenntnis*. Retrieved
December 18, 2018, doi.org/10.1007/s10670-018-0079-4

Cohen, G.A. (2008). *Rescuing Justice and Equality*. (Cambridge, Massachusetts: Harvard
University Press)

Nussbaum, M.C. (2000). *Women and Human Development*. (Cambridge, United Kingdom:
Cambridge University Press)

Rawls, J. (1996). *Political Liberalism*. Paperback Edition. (New York, New York: Columbia
University Press)

Rawls, J. (1999). *A Theory of Justice.* Revised Edition. (Cambridge, Massachusetts: The Belknap
Press of Harvard University Press)

Rawls, J. (2001). *Justice as Fairness: A Restatement*. (Cambridge, Massachusetts: The Belknap
Press of Harvard University Press)