

**Rawls, Harsanyi and the Veil of Ignorance**  
A defence of the maximin decision rule in the original  
position  
By Frederik Wild

**Abstract**

To John Rawls, the basic structure of society is the determining factor in the distribution of benefits and burdens, and as such, the primary subject of justice. If citizens were to choose the principles governing this structure – behind Rawls’ veil of ignorance, the original position – Rawls would assume them to choose along the maximin decision rule; maximising the worst possible outcome. John Harsanyi rejects Rawls’ assumption of choosing in accordance with maximin, as to him, this would suggest irrationally high levels of risk-aversion by decision makers. In this paper, I will argue against Harsanyi’s criticism of the maximin principle, mainly by questioning the validity of (subjective) probability calculations in the original position. By further pointing towards a key distinction between risk-aversion and uncertainty-aversion, I suggest that Harsanyi’s premise of extreme risk-aversion loses its force. From there, I conclude that Rawls’ inclination towards the maximin principle in establishing his two principles of justice appears defensible.

**I. Introduction**

In the seminal work *A Theory of Justice*<sup>42</sup>, John Rawls tries to provide an alternative to the prevalent theory of utilitarianism in establishing principles of distributive justice. In what he calls *justice as fairness*, Rawls seeks moral principles on which the basic structure of society should be built on. Drawing upon social contract theory, these principles ought to be worked out by free and impartial members of society. Rawls imagines the members to deliberate behind a *veil of ignorance* – placed in what he calls the original position – and to behave in

---

<sup>42</sup> Rawls, J. (1999). *A Theory of Justice*; Revised Edition. Cambridge, MA: Harvard University Press.

accordance with the *maximin* principle. John Harsanyi<sup>43</sup> criticises the use of maximin, as it seems to invoke 1) paradoxical results, and further, 2) implies seemingly infinite levels of risk-aversion that cannot be claimed to be rational. In this paper, I will address both of these objections while focusing on the latter. I will thereby predominantly evaluate if it is a necessary feature of agents to admit to these extreme levels of risk-aversion in choosing Rawls' proposed principles of justice. I will argue that it is precisely this premise of Harsanyi's argument which does not seem to hold given the circumstances imagined in the original position. To do so, I will explicate John Harsanyi's key premise and its reliance on the use of (*subjective*) *probability measures*. I will then illustrate that these probability measures fail to have grip when accurately depicting the circumstances invoked by Rawls' veil of ignorance. Consequently, the usage of maximin in the original position appears defensible and also seems like an attractive decision criterion again; at least in such circumstances.

The discussion proceeds as follows. To begin, I will provide an outline of John Rawls' concept of distributive justice in part II. In part III, I will explicate Harsanyi's criticism of the maximin principle. In part IV, I will introduce two objections to this criticism and from there argue that the use of maximin behind the veil of ignorance is well-founded after all. I will close the discussion with concluding remarks in part V.

## **II. Justice as Fairness**

The social, political and economic structure of a society is the determining factor in the distribution of benefits and burdens, and as

---

<sup>43</sup> Harsanyi, J. C. (1975). *Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory*. *The American Political Science Review*, 69(2), 594-606.

such, according to Rawls, the ‘primary subject of justice’<sup>44</sup>. This basic structure is comprised of (a) political institutions, providing the legislative, executive as well as the judicative powers of the state, the (b) economic system, regulating the norms of property, trade as well as contractual obligations, and (c) norms which regulate permissible forms of the family.<sup>45</sup> By large, these institutions determine what people can and cannot do and have a severe impact on peoples’ lives. Rawls’ *justice as fairness*, as a principle for social justice, seeks to establish morally best<sup>46</sup> background conditions for these societal structures. Therefore, the regulative principles on which these institutions are based ought to be principles that are fair, or in other words, to which all members of society would reasonably agree.

### *The Original Position*

In developing these basic principles of society, Rawls draws upon social contract theory. Rawls’ conception of the social contract is a rather general one and should thereby serve not solely as an agreement to the form of government, but as the guiding principle to all of societal makeup; the basic structure of society. In his account, Rawls is specifically concerned with the moral principles underpinning this basic framework. Understood as such, Rawls’ contract is a hypothetical situation, a thought experiment, or in other words a tool, to unveil just principles on which the basic structure of a well-ordered society<sup>47</sup> is built.

---

<sup>44</sup> Rawls, 1999, p.7.

<sup>45</sup> Freeman, 2016, p.6.

<sup>46</sup> *Best* is not an absolute notion (nor a truth), but rather the result of the mutual agreement of the parties in the original position.

<sup>47</sup> Freeman, 2016, p.38. f.; ‘a society in which (1) everyone agrees to the same principles of justice; (2) the principles are realised in basic institutions and complied with; (3) reasonable persons are morally motivated to comply’.

Note: Rawls does not attribute a moral sense to the parties. Rather, a reasonableness necessary to participate in social life. Therefore, erecting

Rawls imagines a deliberative process in which the members<sup>48</sup> of society, and their conception of a just state, mutually agree upon fair guiding principles of the basic structure. Note that these principles hold in perpetuity and are not due to renegotiation. To remove biases, Rawls deprives citizens of idiosyncratic information: accordingly, the parties are ignorant of *arbitrary*<sup>49</sup> facts, such as their gender, race, age, personal endowments such as strength, abilities and intelligence, as well as their social status and economic position within society. Construed as such, citizens will promote their favoured principles behind what Rawls' calls the *veil of ignorance*. Consequently, no one will choose in expectance and/or in correspondence to their advantaged or disadvantaged personal circumstances. Rawls coins this deliberative sphere the *original position*. It is important to note that the parties are not deprived of facts entirely. They are still aware of the 'general facts about human society'<sup>50</sup>, as well as undisputed laws and generalisations of economics, political science, biology and so on. After all, they do have to make an informed decision about just principles of a well-ordered society.<sup>51</sup>

To conclude, in justice as fairness, one does not assume the citizens' inclinations as given, and for them to construct a reality which best fulfils them. Rather, the original position deprives citizens of these circumstances to establish a 'framework of rights and opportunities'<sup>52</sup> in

---

background conditions for a beneficial societal order. It is in this sense 'rational' for members to invoke a sense of justice.

<sup>48</sup> Ibid, p.19; the members can also be understood as representatives, in other words '*trustees*' representing a group of individuals and their interests – in trying to establish basic principles of society.

<sup>49</sup> Arbitrary in that they should not play a role in establishing *moral* principles.

<sup>50</sup> Rawls, 1999, p.119; Freeman, 2016, p.12.

<sup>51</sup> Rawls, p.118 f.

<sup>52</sup> Ibid, p.28.

which every individual can equitably pursue their ends; whatever they may come to be.<sup>53</sup>

### *Fundamental Principles of Justice*

Rawls proposes two fundamental principles of justice. To him, these are principles which every deliberator in the original position would eventually come to promote. I will now establish the reasoning Rawls ascribes his members in choosing these principles, and also outline their content.

To Rawls, members in the original position are guided by instrumental rationality<sup>54</sup>: representatives will favour principles that will best advance their notion of the good – whatever it may come to be<sup>55</sup> – given the information and resources available behind the veil of ignorance. In turn, the ends representatives are ultimately trying to pursue when establishing favourable baseline conditions, are referred to by Rawls as *primary social goods*.<sup>56</sup> A baseline level of these goods will allow agents in achieving their notion of the good once placed in society. These primary social goods include basic rights and liberties, income and wealth, self-respect as well as free choice and the opportunities to pursue a wide array of occupations. It is important to note that the parties are interested in extending their notion of the good over the entire span of

---

<sup>53</sup> It can therefore be suggested, as forwarded by Rawls, that the right comes before/has a priority over the good.

<sup>54</sup> Rawls, 1999, p.124.

<sup>55</sup> Freeman, 2016, p.17. f.; to Rawls, members are able formulate a '*conception of their good* or of their primary values and the best way of life to live given their purposes'. Representatives thus want to establish conditions that will enable them in their pursuit of their rational plan of life; whatever that may come to be for them.

<sup>56</sup> Freeman, 2016; Rawls, 1999, p.22. & p.54; these are 'all-purpose social means' necessary for the exercise and the pursuit of 'a wide variety of conceptions of the good'. Further note: the more (primary social goods) the better.

their life. They do not weigh certain periods of their life as being more or less important. Seen as such, they are keen on upholding their interests and their good in the present as well as in the future, essentially trying to establish a ‘rational plan of life’<sup>57</sup>. Further, the parties in the original position are *mutually disinterested*.<sup>58</sup> This does not mean that they do not care about the sentiments of other people, but rather, that the agents in the original position are acting primarily for the interest of themselves and the ‘causes they care for’<sup>59</sup>. It is under these conditions, and by the consent of rational people and their interests, says Rawls, in which fair principles of justice will emerge.

To Rawls, the corollary is that the original position will require all instrumentally rational representatives to promote the following two principles. The first principle, (I) the *principle of equal liberties*, demands that all citizens are treated as equals, and as such, that each member of society is to enjoy a basic minimum of rights and liberties. The second principle (II) demands that the basic structure should account for arbitrary social or economic inequalities; specifically, such that (a) every citizen, regardless of rank or position, is allowed the opportunity to pursue all offices and positions; and (b) that inequalities in the distribution of income and wealth should be to the greatest benefit of the least advantaged members of society.<sup>60</sup> The latter is what Rawls calls the *difference principle*. The difference principle demands that the inequalities within a society shall work to the advantage of the worst-off members of society; a betterment in their situation.<sup>61</sup> Taken together, these principles will suffice to provide an adequate share of primary social goods for all.

---

<sup>57</sup> Rawls, 1999, p.358.

<sup>58</sup> Ibid, p.12; Freeman, 2016, p.19. f.; mutual disinterest does not imply malevolence.

<sup>59</sup> Freeman, 2016, p.18.

<sup>60</sup> Ibid, p.53, p.266.

<sup>61</sup> Ibid, p.56; e.g. by using the gains created by the well-off to provide training/education to less-advantaged.

And as such, these two principles constitute a fair basis to which neither the arbitrarily better-, nor the arbitrarily worse-off citizen could reasonably object to and would therefore accept as guiding principles of a well-ordered society, argues Rawls.

### **III. Choices under Uncertainty**

John Harsanyi suggests that given the constraints imposed by the veil of ignorance, representatives in the original position are effectively trying to make a rational choice under uncertainty. One prominent decision rule for choices under uncertainty – namely *maximin* – is often claimed<sup>62</sup> to underlie Rawls' reasoning in establishing his two principles of justice. From here onwards, the paper will be concerned with assessing the validity of maximin in the original position, mainly by evaluating Harsanyi's critique of it.

#### *The maximin decision rule*

The maximin decision rule is rather simplistic. It focuses on maximising the worst possible outcome – maximising the minimum – and prescribes to choose the alternative in which the worst outcome is superior to the worst outcome of all other alternatives. Representatives in the original position are ignorant of their eventual social and economic position – for what it is worth, they could find themselves in the most unfavourable place in society. Given that, Rawls says that one should choose as 'if his enemy is to assign him his place'.<sup>63</sup> Considering the magnitude of the choice in the original position and the irrevocability of it, Rawls argues that rational actors would behave in as if adhering to the

---

<sup>62</sup> Harsanyi, 1975; Rawls, 1999, p.72. & p.132.

<sup>63</sup> Rawls, 1999, p.133; note that this does not imply one's social position as chosen by a 'malevolent opponent'.

rather prudent<sup>64</sup> maximin principle.<sup>65</sup> Accordingly, representatives would promote principles of justice that demand a minimum endowment of (I) primary social goods for all, and further, (II) maximises primary social goods of the least advantaged.<sup>66</sup>

### *Risk-aversion*

It has been forwarded repeatedly that in order to choose in accordance with the maximin principle, decision makers must exhibit high risk-aversion.<sup>67</sup> Generally, risk-aversion means that agents favour secure prospects over higher valued, but riskier ones. As the representatives in the original position are essentially facing some ‘kind of superchoice’<sup>68</sup>, the assumption of them behaving risk-averse seems sensible, at least at first sight.

### **John Harsanyi’s critique**

One of the most momentous critiques towards Rawls’ use of maximin was established by John Harsanyi. He argues against maximin as the adequate decision rule in the original position. We will now explicate Harsanyi’s critique as follows. First, we will establish his premise that when faced with uncertainty, decision makers are inherently predisposed to employ *subjective probability measures*. Second, we will illustrate some of Harsanyi’s examples in the usage of maximin and thereby indicate that the adoption of maximin might be based on *unrealistic assumptions* and when applied in the original position, could yield

---

<sup>64</sup> Prudent as in focusing on securing goods for the worst-case scenario rather than seeking possibly ‘risky’ gains.

<sup>65</sup> Arrow, 1973; Freeman, 2016; Harsanyi, 1975.

<sup>66</sup> Freeman, 2016, p.45; Rawls, 1999.

<sup>67</sup> Rawls, 1999, p.72. f.

<sup>68</sup> Freeman, 2016, p.34.



*paradoxical results*. This will lead us to Harsanyi's suggestion in using the *equiprobability assumption* and the principle of average utility.

### *Subjective Probability Measures*

According to Harsanyi, one of the few decision rules not reliant on calculating subjective probabilities is indeed the maximin principle. The problem is that when facing uncertainty, argues Harsanyi, decision makers will always resort to some form of probability calculation. *Bayesian expected utility* theory is based precisely on this presumption, in that rational individuals cannot refrain from using probability measures while making decisions under uncertainty. The corollary is that individuals will always try to maximise their expected utility 'computed on the basis of some set of subjective probabilities'<sup>69</sup>. Under Bayes rule, a decision maker assigns probabilities to outcomes with their respective payoffs. From there, the alternative generating the highest expected utility – given by multiplying payoff and probability – should be chosen.<sup>70</sup> Consider the following situation proposed by Harsanyi, in which a player is asked to choose between two bets<sup>71</sup>:

**Bet 1:** *I shall pay you \$100 if candidate X wins the next election, and pay you nothing if he loses*

~

**Bet 2:** *I shall pay you \$100 if candidate X loses the next election, and pay you nothing if he wins*

Depending on which bet is chosen, one is able to assign the

---

<sup>69</sup> Harsanyi, 1975, p.599.

<sup>70</sup> Ibid, 1975, p.599; two of the rationality postulates essential for this argument are: (1) 'If you prefer A to B, and prefer B to C, then consistency requires that you should also prefer A to C'; (2) 'You are better off if you are offered a more valuable prize with a given probability, than if you are offered a less valuable prize with the same probability'.

<sup>71</sup> Ibid, p.599.

player a higher or lower subjective probability to one of the outcomes. For example, if she chose **Bet 1**, her subjective probability of candidate X winning the election must be equal or higher than  $\frac{1}{2}$ . Otherwise she would have chosen **Bet 2**. Agents will always resort to this behaviour when faced with uncertainty, so goes the argument. It follows, argues Harsanyi, that it will be impossible for decision makers to disregard the subconscious assignment of probabilities, even if one does not ‘explicitly think in terms of probabilities’<sup>72</sup>. And this holds even for the maximin case, whereby an agent will implicitly assign a probability of one – or close to one – to the worst possible event, claims Harsanyi.

*Unrealistic assumptions*

Let us assess the maximin principle on an example given by Harsanyi. Imagine that you currently live in New York and have to decide between two job offers. One in New York and one in Chicago. You prefer the one in Chicago. But in order to attend to it, you would have to fly to Chicago tomorrow – which of course involves the small risk of dying in a plane crash. Therefore, your choices are taking a job in Chicago or taking a job in New York, resulting in four possible outcomes, depicted below.<sup>73</sup>

	If the N.Y.–Chicago plane has an accident	If the N.Y.–Chicago plane has no accident
If you choose the N.Y. job	You will have a poor job, but will stay alive	You will have a poor job, but will stay alive
If you choose the Chicago job	You will die	You will have an excellent job and will stay alive

<sup>72</sup> Harsanyi, 1975, p.599.

<sup>73</sup> Example adopted from Harsanyi, 1975, p.595.

The best outcome if you choose *Chicago* is that you will not die in a plane crash while flying there and are therefore enjoying the benefits of the job you like much better. The worst one being the sudden death en route to Chicago. The two outcomes if you choose *New York* are essentially equivalent, as although you have the worse job, you are alive at least in both scenarios – whether the plane crashes or not. Since maximin prescribes to choose the action which maximises the worst possible outcome, you would choose to accept the New York job; given that both outcomes are essentially the same and there is no possibility of dying in either case.<sup>74</sup> In fact, the maximin principle prescribes that you must not choose the Chicago job under any possible circumstance. Regardless of how insignificantly low you estimate the probability of a plane crash, or how minute the chance of it actually occurring is.<sup>75</sup> As long as the probability is greater than zero, maximin will demand it to be the deciding criterion. And if we accept the decision makers' use of subjective probabilities, this can only be rational if you assign a relatively high probability of a plane crash while flying to Chicago, claims Harsanyi.

What Harsanyi hints to is that by applying the maximin decision rule strictly, it might prescribe wholly unintuitive measures; not only in abstract cases such as above, but also in everyday life: one might never again be able to cross a street or drive over a bridge. In both cases, the possible (life-threatening) dangers – i.e. getting hit by a car or that the bridge collapses – must be considered with highest lexical<sup>76</sup> rank; regardless of its infinitesimal chance of ever occurring. To Harsanyi, and to most of us, as manifested in our everyday choices, it is highly irrational to make your behaviour fully contingent on some outcome – and thereby

---

<sup>74</sup> Dying being the obvious worst-case here.

<sup>75</sup> Harsanyi, 1975, p.596.

<sup>76</sup> Lexical as adapted from Harsanyi; meaning that the danger has to be ranked lexicographically higher according to maximin; and therefore assessed prior to the possible benefit of the outcome.

show an extreme aversion to risk – no matter how unrealistic its occurrence is.

### *Paradoxical Results*

Harsanyi provides us with a further example of how maximin fares; this time when applied to decisions relevant to the original position. Imagine a society of two people, A and B. There is a surplus of an educational grant. The grant is indivisible and so either A or B will receive the fund in whole. Whereas A is a bright scholar and could derive immense satisfaction from the improved educational resources, B is cognitively impaired and would only marginally benefit from them, bettering his life circumstances only slightly.<sup>77</sup> One of the implications of the original position is that the parties will consider the possibility of inheriting the position of the least-favoured citizen, manifested by the *difference principle*. To Harsanyi, the maximin decision rule – and the difference principle accordingly – would then prescribe the grant to be given to B, as it requires to maximise the conditions of the citizen in the worst position. Harsanyi claims this result paradoxical, as the grant would generate a much higher utility and much ‘more intensive human satisfaction’<sup>78</sup> for A. To him, both a utilitarian approach and one’s intuition would require the fund to be spent on A.

### *Equiprobability Assumption*

Reasoning based on the maximin principle would suggest representatives placing an extremely large probability – maybe even unitary/one – on the least-favourable outcome, claims Harsanyi. To him,

---

<sup>77</sup> Ibid p.596; the effect on B’s condition: only minor improvements such as helping him learn ‘how tie his shoelaces’.

<sup>78</sup> Ibid, p.596.

this extremely risk-averse behaviour of assigning exceedingly high probabilities to the possibility of being the worst-off, or at least one of the worse-off in society, is extremely irrational. It simply overstates the possibility of these worst-case scenarios ever occurring. Harsanyi even claims it to be ‘foolish behaviour’<sup>79</sup>. Further, imagine the original position a historical fact.<sup>80</sup> Any member of society not belonging to the worst-off might then have a ‘legitimate complaint against his ancestor’<sup>81</sup> in choosing Rawls’ maximin-based principles. Had the ancestor chosen less risk-averse and more rational so to speak, one could have been made better off. After all, citizens would only benefit from risk-averse decisions of their ancestors, if they indeed find themselves as members of the poorer groups, if not poorest group of society, says Harsanyi.

Given all this, Harsanyi forwards that when faced with uncertainty – not knowing which place in society one will take – a more convincing course of action in the original position would be to assign equal probabilities to being the best-off, the second best-off, the third best off and so on, all the way down to the worst-off member of society. Therefore, the probability given  $n$  individuals would be  $\frac{1}{n}$ . This describes the *equiprobability assumption*. In this fashion, an individual would choose principles of justice based on the highest (expected) average utility level. Compared to maximin, this method would be easier to justify to all members of the society, at least claims Harsanyi.

#### **IV. Objections**

I will now establish two objections towards Harsanyi’s critique. The first is aimed at his argument of paradoxical results. The second

---

<sup>79</sup> Harsanyi, 1975, p.599.

<sup>80</sup> As in: the deliberative process has already occurred. You are now assessing at it as a particular member of society.

<sup>81</sup> Harsanyi, 1975; p.599.

targets Harsanyi's premise that agents are able to develop and employ (subjective) probability measures behind the veil of ignorance.

### *First Objection*

Recall Harsanyi's example of the educational grant and the paradoxical result the maximin principle seems to invoke. First off, I disagree with the view that Rawls conceives his established principles – especially the difference principle – as emanating solely from the maximin decision criterion and for them to be simplistically equated with it. Moreover, specifically the difference principle is not imagined to be applied to such individualistic cases, much less when concerned with the allotment of goods to specific persons.<sup>82</sup> Rawls establishes moral principles underlying society's institutional arrangements, not as a ready-made guidance to small-scale cases such as the one sketched out by Harsanyi. Secondly, his two principles of justice taken together work in a way that would prevent such rather abstract situations from occurring, particularly accompanied by the societal precondition Rawls calls *close-knitnes*<sup>83</sup>: To Rawls, society is inherently connected: making the more advantaged better off will also mean to make the less advantaged better off, and for that matter, all agents in between. Understood as such, we could imagine a scenario in which Rawls' principles would indeed allow the educational funds to be given to A, while assuming that the larger utility A could create with it will sequentially benefit the less well-off people within society; eventually creating better circumstances for B also.<sup>84</sup> Therefore, Rawls' societal order would not demand the funds to

---

<sup>82</sup> Rawls, 1999, p.56; and especially not to these rather special cases ~ disabled persons and the like.

<sup>83</sup> Rawls, 1999, p.70-72; or 'chain-connection'; to Arrow, 1973; p.252; '*close-knitnedness*'.

<sup>84</sup> Some have equated Rawls' difference principle with trickle-down economics: assuming that benefits of the wealthy will favour the society at

be given to B unequivocally, like Harsanyi claims. Of course, in practice, the assumption of a close-knitted society is a strong one to make. By looking at empirical examples, one may point to instances<sup>85</sup> in which the wealthy benefit, and the poor do not; or even worse, where the wealthy benefit at the expense of the poor. This second part of the objection might therefore be stained – albeit empirically – and Harsanyi can be said to have unveiled some shortcomings of the maximin principle; at least when applied to these rather abstract circumstances.

### *Second Objection*

The second objection tries to undermine Harsanyi's use of expected utility theory in the original position. I will argue that representatives behind the veil of ignorance do not seem to have sufficient basis for formulating the probabilities that are necessary for Harsanyi in calculating expected utility as well as declaring maximin as irrational.<sup>86</sup> Even further, I argue that we might not even warrant agents to employ Harsanyi's rather simplistic equiprobability assumption. I will illustrate my argument on the following example.<sup>87</sup>

Imagine a situation resembling that of the original position: you are the class speaker and have to decide on a distribution of grades available for all 20 students in your class. None of them (including you) knows which student one will come to be and as such, which grade<sup>88</sup> one will receive. You are proposed two distributions: on the first, there are

---

large, and therefore also the poor(er). Less strong, Rawls merely assumes that the wealth creation by the well-off will suffice to increase the circumstances of the less well-off; by providing training or similar.

<sup>85</sup> As in Arrow, 1973, p.252; a simultaneous reduction of income tax for high brackets and welfare payments. Research by the IMF (Norris, et al., 2015) suggests: trickle-down does not work in practice, at least from top to bottom. In contrast, 'an increase in the income share of the bottom 20 percent is associated with higher GDP growth'.

<sup>86</sup> Rawls, 1999, p.150.

<sup>87</sup> Example adopted from Schroeder, 2007, p.2.

<sup>88</sup> And thereby, what utility from that grade one will derive.

18 As and two Ds. On the second, there are simply 20 D+s. Which pool of grades would you choose if you are trying to establish best possible conditions for *your* future grade allotment? At first sight, the first seems to be the sensible choice. For one, the chances of getting an A are relatively high and two, the equiprobability assumption, leading us to maximise the average grade, also clearly favours the first distribution: assuming equal chances, one will receive an average utility of  $(18As+2Ds)/n$  where  $n = 20$ . Clearly a higher average utility than  $20D+s/20$ . Given this, Harsanyi would argue for the first distribution of grades. However, what if you assume, as does Rawls, that the final choice of (grade) allocation – once the distribution is set – will be made by your worst enemy?<sup>89</sup> In this respect, assuming the worst, would leave you with a D in the first distribution of grades and a D+ in the second distribution. The second distribution now seems to be the smarter choice in maximising your personal (grade) outcome.

The point of this example is to suggest that the veil of ignorance, as construed by Rawls, renders it impossible to calculate expected utility measures, let alone reliable ones. Such calculations would require one to have at hand at least some well-grounded basis on how one's place in society will eventually be assigned, and also, which of these workings is the more likely: be it by mere randomness, the equiprobability assumption, your worst enemy, picking last, or in any other thinkable way. The problem is that behind the veil of ignorance, you are not aware of any such evidence that would allow you to form the probabilities necessary for utility measures. Even less so, you do not even have historical knowledge of your society, nor do you have information on the possible social and economic states of your future society; contrary to what numerical/ distributional examples – like our classroom case –

---

<sup>89</sup> Rawls, 1999, p.133; Schroeder, 2007, p.2.; or e.g. that you will get to pick last.



might suggest.<sup>90</sup> All these ambiguities will ultimately leave one paralysed when faced with life-determining (distributional) decisions such as the one imagined in the original position.

*Rethinking the original position*

At this point, one might call attention to the distinction between risk-aversion and uncertainty-aversion.<sup>91</sup> Being averse to a known gamble is different than being averse to an entirely obscure or ambiguous situation; to say, in which one cannot come to know the odds. Risk-averse people do not like to gamble when the odds *are* known. Uncertainty-averse people do not like gambles when the odds are *not* known.

Harsanyi's alleged implication of infinite risk-aversion behind the veil of ignorance might thereby be circumvented. As demonstrated, representatives do not seem to be able to form subjective probabilities; at least on reasonable grounds.<sup>92</sup> In other words, members in the original position do not have justified belief in how the odds are distributed. To say, Rawls' proposed veil is not one of mere risk, but rather, one of complete ignorance: of *unmeasurable uncertainty*.<sup>93</sup> This, I argue, is precisely what Rawls aims to achieve with the veil of ignorance. You do not know where you will end up, how these places might look like, and most

---

<sup>90</sup> Rawls, 1999, p.134 f. Think of Harsanyi's examples. One might claim that maximin is only irrational, because agents have empirical knowledge about plane crashes and car accidents – and their extreme unlikeliness. Further, agents have knowledge about the 'workings' of these instances, contrary to members of the original position.

<sup>91</sup> Among others, suggested by Susan Hurley in *Natural Reasons* (1992), p.374 f. A result emanating from the Ellsberg Paradox. Also called ambiguity-aversion.

<sup>92</sup> Hurley, 1992, p. 375; at most, they might form 'unreliable probability judgements, based on ignorance'.

<sup>93</sup> This term stems from Frank Knight's *Risk, Uncertainty, and Profit* (1921) p.20. Distinguishes measurable from unmeasurable uncertainty. In the latter 'risk is not susceptible to measurement'.

importantly, along which rule this societal position is going to be assigned. And even by applying the equiprobability measure – thereby accepting unawareness of reliable odds – an agent will somewhat naively assume equal chances of ending up in any given societal position. But how can one legitimately make such an assumption? After all, representatives in the original position have no evidence on the workings and the likelihood which place one will take. Hence, assuming equal chances seems to me like a rather ill-founded response to the uncertainty envisioned behind the veil of ignorance.<sup>94</sup>

If this objection appears convincing, agents cannot be claimed to reasonably employ their subjective probability measures, nor rely on the rather simplistic equiprobability assumption when placed behind the veil of ignorance. I would then argue that wholly deprived of probabilities, Harsanyi's expected utility calculations do not seem to have grip, and further, that his presumed disposition of high risk-aversion is not a necessary condition for Rawls' representatives.<sup>95</sup> Given all this, deciding in accordance with maximin and forwarding Rawls' two fundamental principles of distributive justice – thereby securing one's fundamental share of primary social goods – again appears like a rational response by uncertainty-averse agents trying to minimise their (unmeasurable) uncertainty.

## **V. Conclusion**

To Harsanyi, representatives behind the veil of ignorance would behave exceedingly irrational by choosing in accordance with the

---

<sup>94</sup> Rawls, 1999, p.152; 'there are no objective grounds in the original position for accepting equal likelihoods, or indeed, any other probability distribution. These likelihoods are merely as-if probabilities. They depend solely on the principle of insufficient reason'.

<sup>95</sup> Rawls, 1999, p.118; especially as the knowledge of one's attitudes towards risk is blocked by the veil of ignorance.

maximin principle. As outlined in this paper, by adhering to his argument (of infinite risk-aversion) in judging maximin and its implications as irrational, one fails to accurately depict the circumstances of Rawls' original position.

Of course, some might still not be persuaded that every representative in the original position would choose in line with maximin 'come what may'. One might argue that in the end, people do exhibit the tendency to engage in gambles<sup>96</sup> even when confronted with unmeasurable uncertainty<sup>97</sup>. We might even agree in that no one can absolve oneself of such a kink<sup>98</sup>. However, with this paper, I hope to have shown that Rawls' original position is a sphere where such gambles are neither well-advised nor well-founded. After all, we are trying to set the most basic moral principles for the society in which we – and our descendants – inevitably have to live in. It might not be smart to gamble away one's basic liberties, opportunities or resources just for the chance of – maybe only marginally – greater endowments. Thus, I agree with Rawls in that representatives will 'prefer to secure their liberties straightway rather than have them depend upon what may be uncertain and *speculative actuarial calculations*<sup>99</sup>. And I would therefore also make the case that representatives in the original position will eventually come to promote his two principles of justice.

## **Bibliography**

---

<sup>96</sup> Schroeder, 2007; for example, a guaranteed Euro vs a 99% chance at one million Euros. Seems like a 'gamble' everyone would engage in.

<sup>97</sup> Knights' unmeasurable uncertainty; also called 'Knightean uncertainty'.

<sup>98</sup> Kink refers to gambling here.

<sup>99</sup> Rawls, 1999, p.139. italics added.

Arrow, K. J. (1973, May). Some Ordinalist-Utilitarian Notes on Rawls's Theory of Justice by John Rawls. *The Journal of Philosophy*, 70(9), 245-263.

Dabla-Norris, E., Kochhar, K., Suphaphiphat, N., Ricka, F., & Tsounta, E. (2015). *Causes and Consequences of Income Inequality : A Global Perspective*. International Monetary Fund. International Monetary Fund.

Freeman, S. (2016). *Original Position*. (E. N. Zalta, Ed.) Retrieved from The Stanford Encyclopedia of Philosophy : <https://plato.stanford.edu/archives/win2016/entries/original-position/>

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

Hurley, S. (1992). *Natural Reasons: Personality and Polity*. New York: Oxford University Press.

Knight, F. H. (1921). *Risk, Uncertainty, and Profit*. Boston, MA: Hart, Schaffner & Marx; Houghton Mifflin Co.

Rawls, J. (1999). *A Theory of Justice; Revised Edition* (Revised Edition ed.). Cambridge, MA: Harvard University Press.

Schroeder, A. (2007). *Rawls and Risk Aversion*. Retrieved from [http://www1.cmc.edu/pages/faculty/ASchroeder/docs/Rawls\\_Maximin.pdf](http://www1.cmc.edu/pages/faculty/ASchroeder/docs/Rawls_Maximin.pdf)