Economics & Philosophy (2025), 1–20
doi:10.1017/S0266267124000452

REPLY

Avoiding risks behind the veil of ignorance

Paul Weithman

406 Malloy Hall, Department of Philosophy, University of Notre Dame, Notre Dame, IN 46556, USA
Email: pweithma@nd.edu

(Received 31 May 2024; revised 17 October 2024; accepted 24 October 2024)

Abstract

Lara Buchak defends a Weight-Ranked Utilitarianism (WRU) that she says avoids the critique of Rawls's that is sometimes thought fatal: utilitarianism unjustifiably blurs the distinction between persons. Buchak's defence depends upon (i) a version of Harsanyi's assumption that parties to a social contract should reason as if they have an equal chance of being anyone and (ii) a hypothesis she explores in a recent article. I argue that her assumption and hypothesis are untenable. WRU fails of the generality to which Buchak aspires because it fails for one of her most important cases: the distributive question posed by Rawls.

Keywords: Lara Buchak; equiprobability; John Harsanyi; weight-ranked utilitarianism; rank-weighted expected utility

John Rawls famously criticized utilitarianism for its failure to take seriously the distinction between persons (Rawls 1999a: 24). The ground of this criticism is that in utilitarianism, Rawls said, 'the principle of rational choice for one man is taken as the principle of social choice as well' (Rawls 1999a: 263). More specifically, utilitarianism mistakenly – Rawls thought – holds that the principles for distributing benefits over stages of an individual's life should also guide the distribution of benefits across persons.

Lara Buchak has recently defended a version of utilitarianism that she says meets this criticism: Weight-Ranked Utilitarianism – hereafter 'WRU' (Buchak 2017). This is a form of utilitarianism that attaches different weights to the interests of different individuals, with the weights depending upon where individuals are positioned or ranked in a particular distribution (Buchak 2017: 620). Buchak, like Rawls, selects distributive principles using an original position: a social contract situation in which contracting parties are veiled in ignorance. But in Rawls's original position, parties choose the difference principle after asking how they would fare if they were in the least advantaged social position. In Buchak's, the parties choose WRU after taking into account the prospects of their being in *any* social position, as well as their attitudes toward risky prospects. Like the utilitarians Rawls criticizes, Buchak extends the principle of choice for one person to a principle of social distribution. As I explain in §1, she does so by equating parties' attitude toward risk within a life with the importance or priority accorded to positions in the social distribution. She claims to avoid the counterintuitive implications that led Rawls to criticize such an extension by requiring that parties' attitude toward risk be reasonable, so that the priority or 'importance function' with which the attitude is equated is also reasonable.

Buchak intends WRU to apply to a broad range of distributive questions, including the distributive question addressed by Rawls: the question of what principles should govern the basic structure's distribution of primary goods. Buchak's work on distributive ethics has drawn a great deal of critical attention (Thoma and Weinberg 2017; Nebel 2020; van Fossen 2024). To my knowledge, none of Buchak's critics has targeted an assumption she takes over from the economist John Harsanyi: the assumption Harsanyi called 'equiprobability'. That is the assumption that parties tasked with choosing distributive principles should decide as if they have an equal chance of – literally – being anyone. In §2, I show why some such assumption is critical to Buchak's argument for WRU. But beginning in §3, I contend that her argument fails in the case of Rawls's distributive question, and fails because of the untenability of that assumption.

My argument depends upon the fact, established in §3, that Buchak needs a specific version of the equiprobability assumption if she is to defend WRU as the answer to Rawls's distributive question. That version requires that parties in her original position know how the distributive principles they consider will affect their society's class structure. More specifically, it requires them to know, for each distributive principle they consider, how many people would occupy each of the social classes that would result if that principle were adopted.

Buchak's version of equiprobability is in one way weaker than Harsanyi's and may initially seem more plausible. But the requirement that parties know class sizes still makes her version informationally demanding. In §4, I argue that it is implausible to think the information it demands can be had for some of the distributive principles the parties have to consider. Buchak has anticipated and attempted to rebut this line of criticism in a recent article (Buchak 2023). There, she invites us to assume, at least for the sake of argument, that that information *could* be had. She argues that parties in the original position should be able to draw on it because knowing it would not compromise their impartiality. If Buchak is right – if the information could be had and parties could take account of it – then her version of equiprobability would be vindicated after all.

This more recent argument of Buchak's, which I discuss in §5 and §6, is the object of much of my attention. While my goal is neither to interpret nor to defend Rawls, I proceed within a Rawlsian framework and my critique of Buchak turns on a claim that is fundamental to Rawls's critique of utilitarianism. That is his claim that 'the correct regulative principle for a thing depends on the nature of that thing' (Rawls 1999a: 25). The thing to be regulated by principles that answer Rawls's distributive question is the basic structure of society. Because of the nature of that thing, the principles that regulate it are fundamental or first principles of justice. I argue that if the principles chosen in it are to be fundamental, the original position must incorporate moral ideals which are more demanding than the ideal of impartiality that Buchak builds into her version of it. Those ideals are incompatible with allowing parties the information Buchak's version of equiprobability requires them to have. I conclude that while WRU may well provide a 'framework for relating risk and equality' (Buchak 2017: 630) that is appropriate for some distributive questions, it fails of the full generality to which I believe Buchak aspires. It does so because it fails for what I believe she takes to be one of her most important cases: the distributive question Rawls poses.

1. Rank-Weighted Expected Utility and the Risk Principle

I said that Buchak thinks parties in the original position would adopt WRU because they take account of the risks to which their choice exposes them. They do so by calculating what Buchak calls their 'Rank-Weighted Expected Utility' – hereafter 'REU'. She argues that parties should maximize their REU and that doing so will lead them to adopt WRU. Buchak's incorporation of risk calculations into the original position is one of her pioneering contributions. But because my worries about Buchak's argument do not concern the details of REU or the consequences of trying to maximize it, my explanation of REU and WRU will be brief and rudimentary.

REU can best be understood in contrast to expected utility – hereafter 'EU'. EU theory promises guidance to an agent who must decide among a set of actions, $\{A_1 \ldots A_n\}$, each of which has several possible outcomes. The EU of any of one of those actions A_i is, as its name suggests, a measure of how well someone can expect to do by performing it. According to the usual way of explaining EU, the agent is to assign a probability, and a utility or payoff, to each of the outcomes which might eventuate from her performance of A_i. The probability assigned to an outcome of A_i is the agent's likelihood of getting exactly that outcome's payoff. The EU of A_i is arrived at by multiplying each payoff by its probability, and summing the products. EU theory tells the agent to choose the member of $\{A_1 \ldots A_n\}$ with the highest EU.

But there is another method for calculating the EU of actions that yields the same results. The alternate way of calculating starts with A_i's lowest possible payoff, the payoff of the worst of that action's possible outcomes. The agent can expect at least that payoff if she performs the action. But she might reasonably expect to do better than that, perhaps much better, depending upon the likelihood of other, more favourable outcomes. The second way of determining how well she can expect to do assigns probabilities to various gains over the minimum that she might realize. So unlike the first way of calculating EU, the second does not assign probabilities to someone's getting *exactly* each of the possible payoffs. It assigns probabilities to her getting *at least* each of those payoffs. It arrives at the EU of A_i by multiplying those probabilities by the differences in the payoffs above the minimum that the agent could receive by performing A_i and adding them to the value of the payoff she is sure to receive: the minimum. This way of calculating the EU of options yields the same results as the first, and gives the same advice: choose the member of $\{A_1 \ldots A_n\}$ which has the highest EU. It just identifies that member via a different arithmetic route.

REU builds on this second way of calculating EU. To see the motivation for it, consider an agent who must select an action from $\{A_1, A_2\}$, each of which has three possible outcomes. A_2 might have a much higher worst-payoff than A_1, but its middling and highest payoffs may be much lower than the middling and highest payoffs of A_1, which we can suppose has a very high maximum. A_1 may offer the highest EU because its maximum payout is so high. But it is not obviously rational to choose it, pass up what A_2 guarantees, and risk getting the low minimum of A_1, all to have a chance of receiving the incremental gains A_1 offers those who luck into the middle and the top. Rationality seems to require that the agent evaluate the risks as well as the EU of her options.

According to REU theory, an agent's attitude toward risk can be expressed as a risk function r. Crudely put, what someone's risk function captures is the way she feels about likelihoods – specifically, about the likelihoods of the gains on the margin that one *might* get by choosing a risky option. The function r therefore takes as its inputs the probabilities of incremental gains above the minimum. REU theory says to pick the option with the highest REU.[1]

Different people may have different attitudes toward risk. Some may have risk functions which imply that the risk of A_1 is worth running while others may not. Risk-averse agents have risk functions – such as Buchak's example of $r(x) = x^2$ (Buchak 2013: 51) – which take the probabilities plugged into them to outputs which are smaller than the probabilities themselves, thus further reducing the value of unlikely outcomes. Extremely risk-averse agents will have risk functions which take all probabilities less than 1 to zero. These agents will attend only to what they can guarantee themselves under each option, and so choose according to the maximin rule. What matters for present purposes is that, on the proposal now in view, rational choosers take account of their attitudes toward risk. REU is said to improve on EU by showing how they can do it.

In the original position, parties choose principles which determine how benefits are to be distributed across society. Each of the principles presented to them has many possible outcomes, and parties do not know how well-off they or others will be if they adopt Rawls's principles, say, or Harsanyi's principle of average utility. Because of their uncertainty, parties in the original position run risks themselves and impose risks on others. When imposing risk on others, Buchak says, agents are to follow what she calls 'the Risk Principle'. That principle says: 'When making a decision for an individual, choose under the assumption that he has the most risk-avoidant attitude within reason unless we know that he has a different risk attitude, in which case, choose using his risk attitude' (Buchak 2017: 632).

Buchak argues that when parties in the original position assign weight to possible outcomes according to the Risk Principle, and maximize their risk-weighted expected utility, they will choose a form of relative prioritarianism.[2] More specifically, they will choose a position that is intermediate between those that Harsanyi and Rawls each argue for. Whereas Harsanyi holds that each individual's well-being matters equally to the evaluation of a social distribution, and Rawls holds

[1] The formula for calculating the REU of an option with an uncertain outcome is:

$$\mathrm{REU}(g) = u(o_1) + r(\Sigma_{i=2}^{n} p(o_i))(u(o_2) - u(o_1)) + r(\Sigma_{i=3}^{n} p(o_i))(u(o_3) - u(o_2)) + \ldots + r(p(o_n)))(u(o_n) - u(o_{n-1}))$$

(Buchak 2013: 53).

[2] In terms of Enflo (2022: 78), a 'personal version' of relative prioritarianism; see also Buchak (2023).

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

that only the well-being of the worst off matters, relative prioritarianism holds that the well-being of the relatively worse off counts for more than that of the relatively better off but that everyone's wellbeing counts for something (Buchak 2017: 611).

According to this distributive rule, Buchak says 'we ought to help the worse off even if we can help them a bit less than the better off – but if we can help them much less, then we ought to help the better off instead' (Buchak 2017: 622). Because WRU attaches different weights to different persons' interests depending upon their place in the distribution, it allows greater weight to be attached to 'the well-being of the relatively worse off'. So Buchak's form of relative prioritarianism receives what she calls its 'formal realization' in WRU (Buchak 2017: 610). Her claims on behalf of relative prioritarianism are therefore claims on behalf of WRU, and her argument for the latter is her argument for the former.

2. The Argument for WRU

At first blush, Buchak's version of prioritarianism has considerable intuitive appeal. The question is whether it can be defended. Buchak helpfully lays out her argument for WRU in premise-and-conclusion form (2017: 638–639). That argument begins with two crucial assumptions and runs via the injunctions to parties in the original position to maximize REU and to follow the Risk Principle. Here is the argument as Buchak presents it:

(1) Individuals in the original position assign common utilities to outcomes they would experience as other individuals.
(2) Individuals in the original position assign equal subjective probability to being each individual.
(3) Individuals in the original position maximize REU.
(4) We should ascribe to individuals in the original position the most risk-avoidant reasonable risk attitude.
(5) Preferences in the original position dictate what our distributive ethical principle should be.
C: We ought to choose policies that maximize weighted-rank utility with respect to the importance attitude I(p) = r'(p), where r'(p) is the most risk-avoidant attitude within reason.

'In other words', Buchak says, 'we ought to give more weight to the interests of the relatively worse off than to those of the relatively better off; specifically, we ought to give them as much weight as we default to giving to relatively worse states in individual decision making' (Buchak 2017: 639).

At the end of the last section, we saw that Buchak takes WRU to be superior to the distributive principles defended by Rawls. Since step (5) says that 'preferences in the original position dictate what our distributive ethical principle should be', she must think that parties in her version of the original position would prefer WRU to Rawls's principles if asked to rank them. Rawls's principles are supposed to answer the question of how the basic structure of society is to distribute the primary goods

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

of rights, liberties, income, wealth, opportunity and the social bases of self-respect. If parties in the original position would prefer WRU to Rawls's principles, Buchak must think that WRU is also an answer to that question. Since parties arrive at WRU by maximizing REU, as enjoined at step (3), they must be concerned with the distribution of utility rather than of primary goods, which Rawls is at pains to distinguish (Rawls 1999a: 78 and 81). But then parties in Buchak's original position would seem to be addressing a different distributive question than parties in Rawls's. In that case, it would seem, they cannot prefer WRU to Rawls's principles as an answer to Rawls's question.

But the appearance of a difference between Buchak's and Rawls's questions is illusory. For though parties in Buchak's original position consider distributive principles whose adoption risks exposing them to various outcomes, Buchak is not committed to characterizing those outcomes in terms of utility, understood as preference-satisfaction or its welfarist consequences. Rather, she says, she 'use[s] the term 'outcomes' . . . to apply to *whatever* is to be distributed' (2017: 625, emphasis added). And while (3)'s reference to REU suggests that what gets distributed is utility, Buchak understands utility capaciously as 'a measure of how valuable each outcome is to the individual for whom it obtains' (2017: 612). Utility as Buchak understands it can be measured using a Rawlsian index of primary goods. (3)'s reference to utility is therefore consistent with Buchak's use of her framework to address Rawls's question of how primary goods are to be distributed by the basic structure.

Once we see how Buchak understands 'outcomes' and 'utility', we can see that when the parties in her original position address that question and maximize REU, what they are to maximize is not their risk-weighted utility as utilitarians typically understand that term. What they are to maximize is their risk-weighted expectation of receiving a certain index of Rawlsian primary goods. And when Buchak concludes that they would choose WRU, what she really thinks is not that the basic structure should maximize some function of utility as utilitarians understand it. She thinks what they would choose is a principle according to which the basic structure should maximize the weight-ranked index.

Though Buchak's original position differs from both Harsanyi's and Rawls's in satisfying (3) and (4), Buchak implies that in her use of the device, she follows Harsanyi more closely than Rawls (Buchak 2017: 611). One of the ways she does so is by taking steps (1) and (2) – both of which she finds in Harsanyi's work, and both of which Rawls explicitly denies. Though Buchak assumes both (1) and (2) without argument, I believe (2) is deeply problematic.[3] So in what follows, I shall accept (5) and will ignore both (3) and (4). While (1) might seem highly contentious – and some versions of it are – I shall grant it as well. For once we see what Buchak means by (2), we will see that the version of (1) she needs is unobjectionable. The difficulties with her argument lie in the second assumption rather than the first, and that is the step on which I shall concentrate.

[3] Buchak (2017: 629) says she follows Harsanyi in thinking (2) is necessary 'for choices in the original position to represent the idea that all people are treated equally'. For an elegant argument that – in effect – rejects (2) on the ground that subjective probabilities need not be equal to secure that result, see Mongin (2001).

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

3. Equiprobability

Some assumption about the probabilities available in the original position is crucial to Buchak's argument. For parties in her original position are to calculate risk-weighted expected utilities, and to do that, they must have probabilities to plug into their risk function. Step (2) – 'the equiprobability assumption' or 'the equiprobability postulate' – provides them the probability assignments they need.

Harsanyi himself phrases the assumption in different ways at different times, ways that are not obviously equivalent. At one place, he explains the postulate by first assuming a society composed of n individuals. These individuals, he then says, are 'numbered as individual 1, 2, . . . , n, according to whether they would occupy the 1st (highest), 2nd (second highest), . . . , nth (lowest) social position under a given social system' (Harsanyi 1982: 45). 'By the equiprobability postulate', he continues,

> individual i [or his representative in the original position] will act in such a way as if he assigned the same probability 1/n to his occupying any particular social position and, therefore, to his utility reaching any one of the utility levels $U_{i,}$ U_2, . . . , U_n . . . (Harsanyi 1982: 45)

In a later formulation, Harsanyi says that each party in the original position 'must assume that he has the *same probability* 1/*n* of ending up in the social position of any individual *i* with *i's* utility function *U;* as his own function' (Harsanyi 1992: 676 (italics original, underlining added)). The underlined portion of this later passage is not included in or implied by the earlier one. It commits Harsanyi to a very strong version of (2), since it requires each of the parties in the original position to imagine what it would be like to be each of the other people in the population, complete with that person's systems of ends.[4] It thereby commits Harsanyi to a very strong version of (1) and it is Harsanyi's versions of (2) and (1) that Rawls objects to in *A Theory of Justice* (Rawls 1999a: 151–152).

I noted above that Rawls's distributive question concerns the distribution of primary goods, so that how well off each individual is is determined by her index or stock of those goods. When Buchak uses the original position to address Rawls's question, all parties in it know that the utilities others attach to the outcomes of her principle and Rawls's are measured by that index. So for purposes of comparing those principles, the version of (1) that Buchak needs is much weaker than the version Harsanyi requires. She needs only

(1') Individuals in the original position all attach the same value to the stock of primary goods they would receive were they to get that stock.

Since parties in the original position assume that everyone attaches the same value to primary goods, (1') is trivially true, and Buchak's assumption of it seems unproblematic. Moreover, because welfare is, for present purposes, measured using primary goods, parties need not know all n utility functions, nor need they imagine

[4] Harsanyi may well have settled on the strong version at the time he wrote (1982), or well before, but his earliest reference to equiprobability in (1953: 435) is to the weak version.

replacing their own utility function with anyone else's. So, again for purposes of comparing her principles with Rawls's, Buchak does not need the strong form of (2) that Harsanyi assumes and that Rawls criticizes, and she does not ever make the assumption in that form. What version of (2) does she need?

Harsanyi expressed the equiprobability postulate in terms of the chances of occupying social positions. So long as someone's social position is determined by her utility as Harsanyi understands it, his assumption that there are as many social positions as there are individuals seems not implausible, since it is not implausible to suppose that no two people have exactly the same level of utility so understood.[5] But once social positions – and the value of occupying them – are given by the index of primary goods, Harsanyi's assumption that the number of social positions equals the number of people seems far less plausible, since many people could have the same index and hence occupy the same social position.[6] So when addressing Rawls's distributive question, perhaps equiprobability should be taken to say – as (2) does – that 'individuals in the original position assign equal subjective probability to being each individual' rather than that they should assign equal subjective probability to occupying each position.

Of course, this version of equiprobability could still be expressed by reference to social positions, defined by indices of primary goods, rather than to individuals. But what equiprobability so expressed would say is not, as Harsanyi has it, that each individual is to take the probability of occupying a social position to be 1/n. Rather, it would say:

> (2') For each of the n social positions that is occupied, individuals in the original position take the probability of their occupying that position to be 1/n multiplied by the number of people who occupy it.

(2') is equivalent to (2). It has the advantage of making clear – in a way that (2) does not – just how parties' REU calculation will go. For it makes immediately evident what probabilities have to be plugged into the risk function and then multiplied by the levels of primary goods that define the utility of each position.

The problem with (2'), and with (2), is that neither is sensitive to important facts about the consequences of addressing Rawls's distributive question in the original position. Those are that what social positions will be occupied and how many people will occupy them depend upon which principles parties in the original position adopt and how the adopted principles are implemented at subsequent stages of what Rawls calls 'the four-stage sequence' (Rawls 1999a: 171–176). If the parties choose a

[5] Here I assume that the function which takes resources to utility is real-valued. Whether Harsanyi understood 'utility' in the usual utilitarian fashion is the subject of some dispute; see Sen (1986: 1122–1124) and Risse (2002). For textual evidence that he does, see Weymark (1991: 308).

[6] Harsanyi (1953: 434–435) took alternative social states to be alternative distributions of income and took a person's social position to be given by her place in the distribution, so in that article he did not assume that social positions and population were equinumerous. As John Weymark (1995: 314) observes, it is only in Harsanyi's later work that he changed his view so that 'social alternatives are permitted to include as many factors as are considered relevant'. Once this permission is granted, the assumption of equinumerosity seems more plausible. And as Weymark adds immediately, it was only granting this permission that 'Harsanyi was able to link Welfare Economics with expected utility theory'.

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

conception that is implemented in loosely regulated markets and in a light-handed tax regime, the post-tax social classes and the number who occupy them would be quite different than if they choose a conception that requires a steeply progressive tax structure.

When evaluating the options before them, parties in the original position need enough information to consider how each of the options will affect those in the social positions deemed relevant to the parties' choice. One of the attractions of maximin is its informational simplicity: parties who employ that rule need only consider the expectations of those who are worst-off under each option. Buchak's rejection of maximin means that she needs a version of (2') which is sensitive to the effects of parties' choices on all social positions. And so I believe that instead of (2'), what she needs is the informationally demanding:

(2") For each conception of justice C that individuals in the original position consider:

- those individuals have to know what social positions will be occupied if they choose C and how many people will occupy them, and
- for each of the n social positions that is occupied in C, they take the probability of their occupying that position to be 1/n multiplied by the number of people who occupy it.

But if (2") is the version of (2) that Buchak needs, then I think her equiprobability assumption cannot be right.

4. A Doubt about (2")

Parties in Rawls's original position decide according to the maximin rule because Rawls denies them information on the basis of which to assign likelihoods, and so denies them the information they would have to have for (2") to be true of them. Their lack of information results from the veil of ignorance. Buchak's disagreement with Rawls might therefore be thought a disagreement about what information should be allowed through the veil – a disagreement about, as Rawls might put it, how thick or thin the veil of ignorance should be (Rawls 1999b: 336). I shall consider this possibility in the next section. First, I want to look briefly at other reasons for denying (2").

One of the conceptions of justice among which parties in the original position have to decide is average utilitarianism, the conception that Rawls takes to be the primary rival of justice as fairness and that Harsanyi thought parties would choose. Buchak wants to add WRU to the list of conceptions they consider. If 'utility' in WRU is reckoned in terms of primary goods, then if parties adopt it, primary goods will be distributed so as to maximize weight-ranked utility so understood. But I shall assume that average utilitarianism is committed to the standard utilitarian way of understanding utility. If parties adopt *it*, primary goods will be distributed so as to maximize average utility as utilitarians typically understand it. According to the utilitarianism with which Rawls contrasted his own view, utility was understood as the satisfaction of rational desire (Rawls 1999a: 22–23). But for present purposes the

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

utility that average utilitarianism says to maximize can be understood, as I believe Harsanyi understood it, as welfare, without commitment to the thesis that welfare is engendered by desire-satisfaction. To decide between average utilitarianism and justice as fairness – or between average utilitarianism and WRU – parties have to consider the implications of distributing primary goods as average utilitarianism dictates. Can the probabilities referred to in (2″) guide their decision-making?

If the parties are to know what they would need to know for (2″) to be true of them, they would have to know what social positions will be occupied and how many people will occupy each position when average utility is maximized. Grant, what is by no means obvious, that maximizing average utility requires equal rights and liberties, and fair equality of opportunity, so that social positions differ only in the remaining primary goods: income, wealth and the bases of self-respect. Now assume, as seems plausible, that there are certain occupations which are socially necessary – firefighter, law enforcement officer, physician, high school teacher, government official, and many more – and that average utility will be maximized only if these are filled.

We can suppose for the moment – and for the sake of argument – that parties in the original position have fine-grained knowledge of their society's history and economy, of what occupations are socially necessary and of how many people need to hold such jobs for average utility to be maximized. Even so, knowing what wages have to be attached to these positions to fill them in a free labour market would require the parties to have quite detailed knowledge of individuals' utility functions. This is because different people might have different work/leisure trade-offs and will derive different amounts of satisfaction from jobs at given wage-levels. It is also because people's satisfaction is interactive: how much utility someone has depends heavily on what she knows about what other people have, whether they are in her circle or at some social distance, whether she needs to provide for them and whether she herself is prone to envy.

It may be said that free labour markets will adjust wages to need, and so ensure that necessary occupations will be filled at the wages workers demand for filling them. If workers do not derive enough utility from market wages to make working in necessary occupations worth their while, they will raise their demands until market wages suffice. But even if labour markets work perfectly, there is no guarantee that market equilibria can be predicted, that markets will reach the same equilibrium point in different societies or that the equilibria they attain will be ones at which average utility will be maximized. Moreover, labour markets do not work perfectly; asymmetries of power and information can result in equilibria which are sub-optimal by any measure, including the measure enjoined by average utilitarianism.

I conjecture – though I cannot prove – that because of these complexities, the information parties in the original position would have to have for (2″) to be true of them is far too detailed and contingent to be knowable even with all of the information I have temporarily supposed is allowed them. In short, it is too detailed and contingent for parties to know what social positions will be occupied should they choose average utilitarianism and how many people will occupy them. If this is right, then once we see what equiprobability assumption Buchak needs to address

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

Rawls's distributive question in her version of the original position – an assumption equivalent to (2'') – we can see that that assumption should be rejected.

5. But What if We *Could* Know?

In a more recent article, Buchak has, in effect, anticipated and responded to the kind of argument against equiprobability that I have just given. She writes:

> The original position is supposed to abstract away from morally irrelevant facts, as well as facts that cannot be known in advance. Thus, if we assume we [i.e. parties in the original position] are ignorant of the probabilities, it must be that the sizes of the social classes under each institutional arrangement are either morally irrelevant or empirically unknowable. (Buchak 2023: 230)

In the last section, I claimed that when parties in the original position consider average utilitarianism, they cannot assign probabilities to social positions because they cannot predict the occupational choices that would be made by members of a society well-ordered by that conception. So I have, in effect, contended that they are ignorant of what they need to know to consider average utilitarianism because 'the sizes of the social classes under [that] institutional arrangement are … empirically unknowable'. But, Buchak says, 'we can ask what would follow if, contrary to supposition, we could know them'. 'What would follow', she asks 'if we discovered some general principles about human behavior that implied that societies with arrangements like A always result in 60% of people ending up in the upper-class and 20% in each of the other two classes?' (Buchak 2023: 231).

Buchak contends that what would follow from that – and so what would follow if information about class size is morally relevant but unknowable – are consequences untoward enough that the reason for excluding that information from the original position must be its moral *ir*relevance. In a passage worth quoting at length, Buchak then distinguishes two ways in which information can be morally irrelevant, and argues that even if information about class sizes *is* morally irrelevant, it should be allowed into the original position. She writes:

> [T]here are … two ways for the size of the social classes to be morally irrelevant. It might be morally irrelevant in the sense that it could cloud my judgment or make my rational interests cease to coincide with what is moral (this is the sense in which the social position I occupy in the actual world is morally irrelevant). Or it might be morally irrelevant in the sense that it is inert: even if I know the size of the social classes, this fact will not make a difference to my judgment, given the decision rule I use. Facts that are morally irrelevant in the former sense should be excluded from the original position; facts that are morally irrelevant in the latter sense should be included in the original position, but should be excluded from making a difference by the decision rule itself.
>
> If the size of the social classes is morally irrelevant, it is clearly morally irrelevant in the latter sense: it is a deliverance of our evaluations of social

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

> gambles, not an assumption about which social gambles make our rational preferences mirror the moral ones. It is not the kind of thing that could cloud my judgment because it makes me favour my own interests, for example, or the interests of those like me; thus, even if it ends up not mattering to the ultimate evaluation of social arrangements, it should not be excluded from the original position (Buchak 2023: 231–232).

I shall not work through this passage sentence by sentence, but I believe Buchak's argument can be laid out as follows:

(1_{CS}) Suppose that information about class sizes under various institutional arrangements is morally irrelevant.

(2_{CS}) There are two and just two ways for information to be morally irrelevant: (i) 'it could cloud my judgment or make my rational interests cease to coincide with what is moral' and (ii) it can be inert.

(3_{CS}) Information about the sizes of social classes under various institutional arrangements 'is not the kind of thing that could cloud my judgment because it makes me favour my own interests, for example, or the interests of those like me'.

(4_{CS}) If information about the sizes of social classes under various institutional arrangements is morally irrelevant, it is morally irrelevant in the second way and not the first.

(5_{CS}) 'Facts that are morally irrelevant in the [first] sense should be excluded from the original position; facts that are morally irrelevant in the [second] sense should be included in the original position[.]'

C_{CS}: Even if information about sizes of social classes under various institutional is morally irrelevant, it should not be excluded from the original position.

Buchak's point is not that we have discovered information about class sizes under average utilitarianism that can be provided to parties in the original position so that (2'') is true of them after all. Her point, rather, is this. Rawls's rationale for setting up the original position so that it would exclude the information about class size on which (2'') would draw – were that information available – depends upon taking that information be morally irrelevant in a sense of 'morally irrelevant' that does not apply to it. Rawls should therefore allow that information into the original position. Thus the connection between equiprobability and the argument for C_{CS}.

Buchak is, in effect, asking us to suppose that there are laws of human behaviour which tell us that distributing primary goods so as to maximize average utility will lead to social classes, defined by shares of primary goods, of certain sizes. Assume as before that maximizing average utility requires equal rights and liberties, and fair equality of opportunity. Then the law Buchak asks us to imagine would imply that distributing income and wealth to maximize average utility will lead to, say, 60% of the population being in the upper class and so to each person's having a 0.6 chance of realizing the wealth-and-income difference between the upper class and the

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

middle. The laws we are asked to imagine would thus yield the probability estimates needed to compare average utilitarianism to other conceptions of justice, the estimates referred to in the second clause of (2'').

Those laws would seem to depend upon regularities governing individuals' ability to convert resources into utility. That ability, in turn, would seem to depend heavily on facts that vary from individual to individual, such as facts about physical and mental health, and about whether some people in society have expensive tastes or are liable to develop adaptive preferences. It would also seem to depend heavily on features that vary from society to society. So what preferences citizens develop, which preferences they can satisfy with a middle class income, and how much welfare they derive from the satisfaction of those desires, would all seem to depend on the productivity of various units of labour (which itself depends on society's level of technological development), on consumption opportunities (which also depend on the level of technological development), on available natural resources, on environmental quality, and on society's investment in public goods such as parks, recreation areas, museums, public entertainment, quality education, quality public transport, childcare and health care. They would seem to depend, too, on the extent to which society guarantees citizens' physical security and the rule of law.

The point now is not that dependence upon these factors renders regularities about citizens' ability to convert resources into utility unknowable, though it may. The point, rather, is that regularities which depend upon variable social conditions thereby depend upon social contingencies. If they do, then so too do the laws relating candidate principles to class sizes. And so what the parties in the original position would know if (2'') is true of them are not laws 'about human behaviour that implied that societies with arrangements like A always result in 60% of people ending up in the upper-class and 20% in each of the other two classes'. What they would know are laws that imply that 'societies with arrangements [enjoined by average utilitarianism, for example] will result in 60% of people ending up in the upper-class and 20% in each of the other two classes when some conditions obtain, but in other class sizes under other cultural, economic and technological conditions'.

Buchak would be right to observe, as (3_{CS}) says, that this knowledge 'is not the kind of thing that could cloud my judgment because it makes me favor my own interests, for example, or the interests of those like me'. But the moves from (3_{CS}) to (4_{CS}), and from there to (C_{CS}), require some further, unstated, assumptions.

C_{CS} is a conclusion about what information should be available to parties in the original position. It depends upon (2_{CS}), (3_{CS}) and (4_{CS}), all of which refer to information that would cloud *my* judgement or make *my* judgement deviate from what is moral. The inference from what information affects me to what information parties in the original position should be allowed is not obviously warranted. But parties in the original position are supposed to be our representatives, who need to be veiled in ignorance because they are relevantly similar to us. So if the original position is properly constructed, the information that 'could cloud my judgment or make my rational interests cease to coincide with what is moral' is the same as the information that 'could cloud … the judgment [of parties in the original position] or make [their] rational interests cease to coincide with what is moral'. On the assumption that the two kinds of information are the same, (5_{CS}) can bridge claims about my judgement and C_{CS}.

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

The move from (3_{CS}) to (4_{CS}) depends on a further assumption: that information 'could cloud my judgement or make my rational interests cease to coincide with what is moral' only if it 'could cloud my judgement because it makes me favour my own interests, for example, or the interests of those like me'. Put more simply, it depends on an *Impartiality Assumption*:

> (IA) Information 'could cloud my judgement or make my rational interests cease to coincide with what is moral' only if it compromises my impartiality.

The assumption needed for (5_{CS}) to play its bridging function, together with (IA), implies that information 'could cloud ... the judgement [of parties in the original position] or make [their] rational interests cease to coincide with what is moral' only if it is the kind of information that would compromise my impartiality. But, again if the original position is properly constructed and the parties are relevantly like us, we can make a *Coincidence Assumption*:

> (CA) Information compromises the parties' impartiality just in case it compromises mine.

(IA) and (CA) imply an important claim about parties in the original position:

> (PI) Information could cloud the parties' judgement or make their rational interests cease to coincide with what is moral only if it compromises their impartiality.

Since information about class sizes would not compromise my impartiality, it follows from (PI) that it does not compromise the parties' either. So that information should not be excluded from the original position. Hence C_{CS}.

But the assumptions that I have interpolated all fail to distinguish various subjects of judgement. Parties in Rawls's original position judge of a special case, for they are adopting principles for the basic structure of society. Even if the only information that would 'cloud [the parties'] judgement' in Buchak's sense of that phrase is information that compromises their impartiality (see Buchak 2017: 635), there could be other ways to 'make [their] rational interest cease to coincide with what is moral', at least in that special case. To see this, let us turn first to Rawls's own reasons for excluding more information from the original position than Buchak would and then see what can be said in favour of the exclusion.[7]

6. Autonomy and Irrelevance

Rawls criticizes utilitarianism for mistaking impersonality for impartiality (Rawls 1999a: 166). And he thinks, as Buchak presumably does, that the original position should model the latter rather than the former. But impartiality is not the only moral value Rawls's original position is to incorporate and the Rawlsian veil of

[7] I am grateful to Jonathan Quong for helpful discussion of the points in this paragraph.

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

ignorance is not imposed only to guarantee it.[8] The veil is imposed to ensure that the principles adopted in the original position are principles of right, and impartiality is just one condition principles have to satisfy. For principles to qualify as principles of right, they 'must hold for everyone in virtue of their being moral persons' (Rawls 1999a: 114). Rawls imposes the veil of ignorance to ensure that principles chosen in the original position satisfy this condition as well. It does so by ensuring that what Rawls calls 'the decisive determining element' (Rawls 1999a: 222) of parties' choice in the original position is just their nature as such persons, and not any contingencies about the society that is to be well-ordered by the principles they adopt.

Rawls is especially clear about this in the Dewey Lectures (Rawls 1999b: 336), and he reiterates the insufficiency of impartiality in 'Basic Structure as Subject'. There he says of the veil of ignorance

> we allow in just enough information to make the agreement rational, though still suitably independent from historical, natural and social happenstance. Considerably more information would be compatible with impartiality but a Kantian view seeks more than this. (Rawls 1996: 273)

I have spoken so far of the way parties in the original position address Rawls's distributive question. Since their answer applies to the basic structure, those principles do not apply to natural moral persons, so universality would seem not to be true of them. But while we cannot act from principles chosen to apply to institutions, those principles are to inform our judgements of right. They yield judgements of justice on which we can act.

I have argued that if (2") is true of parties in the original position, so that they can compare distributive principles for the basic structure, they have to have access to society-specific information, and perhaps information about individual utility profiles. That information has to be allowed through the veil. If WRU is chosen in the original position, in preference to average utilitarianism, then the principles chosen will depend on that information. If WRU informs the sense of justice and constrains other principles adopted in the original position – and the choice of principles is conditional on that society-specific information – then principles will not apply just in virtue of being moral persons. They would apply to us as moral persons who are citizens of a particular society.

If the universality condition is an apt one to impose on principles for the basic structure, as I believe it to be, then allowing specific information through the veil compromises the status of the principles chosen. It should therefore be excluded. If it should, then (PI) is mistaken because it 'make[s the parties'] rational interests cease to coincide with what is moral' even if it does not bias their choice or compromise their impartiality. But without (PI), the argument for C_{CS} does not go through. And without C_{CS}, there is no argument for allowing information to which (2") refers through the veil of ignorance, in the event that that information is available.

[8] For detailed discussion of the moral ideals expressed by Rawls's original position, see Moehler (2016).

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

Rawls's description of his view as Kantian may suggest an objection: that in Rawls's hands, universality imports another, far more controversial, condition. For Rawls thinks that when we act from principles of right which satisfy the universality condition, we express our nature as – that is, we live as – moral persons. To live as a moral person is to live as a free person, in a robust sense of free: it is to live autonomously. Autonomy is an ideal that demands more of the original position than impartiality does. The further requirement that principles chosen in the original position must make autonomy possible is what really motivates the thick veil of ignorance and the rejection of (PI). But autonomy is not just a strong, but also a contestable, ideal. The objector might contest it, and insist that the only reason for veiling parties in ignorance is to ensure their impartiality, all with an eye toward holding onto (PI), vindicating (2") and defending WRU.

Moreover, it may be said, even Rawls came to realize that autonomy is a contestable ideal and so he rejected it when he turned to political liberalism. As part of his turn to political liberalism, he allowed that citizens might act from principles of right on the basis of any of a range of reasonable comprehensive doctrines that overlap on those principles. Some of those doctrines reject the claim that to live as a free person is to live autonomously. To insist that principles of right facilitate autonomy is therefore inconsistent with the attempt to frame a conception of justice that can be the object of an overlapping consensus. But to give up on that insistence is to give up on the argument for excluding society-specific information from the original position and for contesting (PI).

I cannot lay out the response in detail here but in brief:

Start with the claim that autonomy is a contestable value whose endorsement is inconsistent with the aims of political liberalism and that Rawls recognized the inconsistency. In *Political Liberalism*, Rawls distinguishes political autonomy from the ethical autonomy of Kant's moral philosophy. He says that citizens of a well-ordered society can realize full political autonomy when they act from a sense of justice informed by principles chosen in the original position. Their full political autonomy depends upon the rational autonomy of the parties (Rawls 1996: 78–79). Put in Buchak's terms, the parties' 'rational interests will coincide with' – or lead them to choose – 'what is moral' only if they are rationally autonomous. And the rational autonomy of the parties depends upon their being moved to secure the highest-order of interests of those they represent, without knowing the particularities of the society to which they belong.

If this reading of the later Rawls is correct, then even after Rawls's turn to political liberalism, he would have maintained that contingent information, such as would be required to know class sizes, is morally irrelevant to the decision to be made in the original position, but not because it would compromise the parties' impartiality. It is irrelevant because to deem it relevant, and to admit it into the original position, would compromise the rational autonomy of parties and the political autonomy that is supposed to be available to members of a well-ordered society. And so the proponent of Rawlsian political liberalism would continue to deny (PI). But without (PI), the argument for C_{CS} does not go through. And without C_{CS}, there is no argument for allowing information to which (2") refers through the veil of ignorance, in the event that that information is available.

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

As I have argued elsewhere, full political autonomy does not just require that someone act from a sense of right informed by principles chosen in the original position. She also has to take the principles to be justified because they would be adopted there (Weithman 2017: 103). When an overlapping consensus obtains, many citizens may take the principles in the overlap to be justified by their comprehensive doctrines rather than by choice in the original position. So those who act on principles of right on the basis of their comprehensive doctrine do not realize political autonomy. But it is nonetheless important that political autonomy be available in a society characterized by an overlapping consensus. That is, it is important that citizens of such a society have access to the justification political liberalism itself offers for principles of right, a justification of them as chosen in the original position with the informational restrictions of the thick veil of ignorance.

For one thing, the idea that comprehensive doctrines are well enough worked out to have clear implications for politics – in the cases of many doctrines, or at least of the people who hold them – seems to me best described as 'aspirational'. Instead, I suspect that a great many people do not see many connections at all between their comprehensive doctrines and their politics, often because their doctrines are not sufficiently well worked out to establish those connections. There may be comprehensive doctrines that do support a conception of justice and whose adherents accept principles of right because of the support. But our relationships to our comprehensive doctrines may not be stable. People change their doctrines, sometimes quite radically. Faith in one's doctrine waxes and wanes and waxes again over the course of life, or certainly can do so. And yet we want people's allegiance to the political conception of justice to be stable through all these changes in their relationship to their doctrines. If it is to be stable, the public culture will have to contain intellectual resources on which people can fall back to sustain their sense of justice. Among the resources will be a presentation of the public conception of justice as following just from citizens' freedom and equality. That resource is available when the conditions of political autonomy are satisfied. Satisfying them requires rejecting (PI).

Of course, there is some artificiality in insisting that principles not be conditional on society-specific information so that those who act on them can act autonomously, as if acting from unconditional principles and acting autonomously were two different things, the former a means to the latter. For to act from unconditional principles just is to act autonomously. What ultimately needs to be defended is the claim that principles for the basic structure should not be conditioned on society-specific information. The real question, then, is what reasons there are for excluding information about features peculiar to one or another society, so as to make autonomy available.

To see what those reasons are, suppose parties' choice of principles for the basic structure did depend upon such information. Suppose, for example, that the parties rejected average utilitarianism because of what they knew about the size of social classes there would be in their society – and perhaps just in their society – were average utilitarianism adopted. Then their decision would conform to a maxim of the form 'If a member of society S, choose WRU in preference to average utilitarianism'. Their decision will, that is, be hypothetical. That is why those the

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

parties represent would not realize political autonomy by acting on the distributive principle they adopt.

At the outset, I quoted Rawls's remark that 'the correct regulative principle for a thing depends on the nature of that thing' (Rawls 1999a: 25). And so, we might infer, the nature of the thing to be regulated determines how deep the principles are which regulate it. Any of the distributive principles now in view – Rawls's own, the principle of average utility and WRU – would, if adopted, apply to the basic structure. The basic structure is the primary subject of justice. Principles that regulate it are therefore fundamental, or first, principles of justice. Principles which are hypothetical, which apply given the hypothesis that they are to be adopted for a particular society, cannot be used to interrogate the features of that society in virtue of which they are chosen. They cannot, for example, be used to interrogate the features in virtue of which a society's population divides into class sizes.

Among the features I cited are facts about a society's culture, about its citizens' ability to convert resources to utility and all the factors on which that ability depends, and about the informational and power asymmetries that affect its labour markets. These features of a society may be just or unjust. But when parties in an original position are assumed to know them – or to know regularities based upon them – they are taken as given, and so are beyond the critical reach of principles chosen in light of them. This means that the principles chosen cannot play the critical role that fundamental moral principles are to play.

Moreover, it would be natural to wonder if the principles chosen on the basis of such contingencies derive their reason-giving force, and so their moral character, from more general or fundamental principles, which are not conditional on features of this or that society. Thus if it turned out that WRU was appropriate for societies in which it would result in certain class sizes, average utilitarianism for other societies and Rawls's principles for still others, it would be natural to wonder whether there are some deeper principles that explain the pairings (cf. Rawls 1999a: 108). What would determine the content of *those* principles is the free and equal nature of participants in a cooperative scheme that is to be democratic. Freedom and equality are among the moral values that the original position models in virtue of its informational restrictions. To give up on those restrictions is to give up what they make possible: categorical distributive principles which are capable of playing a critical role and which apply to the basic structure of liberal democracies just as such. I believe that these are costs political philosophy should be unwilling to pay.

7. Conclusion

Buchak says that one of the aims is 'to provide a framework for relating risk and inequality'. (2017: 630). In §2, we saw that according to that framework

(5) Preferences in the original position dictate what our distributive ethical principle should be.

We have seen that Buchak's version of the original position requires (2''), according to which parties should be allowed information – including society-specific

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

information – on the basis of which they can assign probabilities to outcomes. I believe Buchak wants her framework, and her version of the original position, to apply to a wide variety of distributive questions and to provide a way of evaluating ethical principles which vary greatly in their scope. Society-specific information may well be needed to settle some of those questions. It will, for example, almost surely be needed to arrive at principles for the distribution of scarce medical resources or principles for climate mitigation whose benefits are uncertain and whose costs will be unequally spread. I have not denied the applicability of Buchak's framework to these very important questions. For all I have said, it may be applicable to them.[9] What I have tried to show is that distributive principles for one very important case – that of the basic structure – should not be chosen on the basis of such information. In that case, I have contended, it is a mistake to take risks behind the veil of ignorance.

Acknowledgements. I am grateful to Zach Barnett, Peter de Marneffe, Michael Mohler, Jonathan Quong, Leif Wenar and John Weymark for helpful comments on earlier drafts.

References

Buchak L. 2013. *Risk and Rationality*. Oxford: Oxford University Press.

Buchak L. 2017. Taking risks behind the veil of ignorance. *Ethics* **127**, 610–644.

Buchak L. 2023. Philosophical foundations for worst-case arguments. *Politics, Philosophy and Economics* **22**, 215–242.

Enflo K. 2022. The equivalence of egalitarianism and prioritarianism. *Journal of Ethics and Social Philosophy* **22**, 74–108.

Harsanyi J. 1953. Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy* **61**, 434–435.

Harsanyi J. 1982. Morality and the theory of rational behavior. In *Utilitarianism and Beyond*, ed A. Sen and B. Williams, 39–62. Cambridge: Cambridge University Press.

Harsanyi J. 1992. Game and decision theoretic models in ethics. In *Handbook of Game Theory*, Volume **1**, ed. R.J. Aumann and S. Hart, 671–707. Amsterdam: Elsevier Science Publishers.

Moehler M. 2016. Impartiality, priority and justice: the veil of ignorance reconsidered. *Journal of Social Philosophy* **47**, 350–367.

Mongin P. 2001. The impartial observer theorem of social ethics. *Economics and Philosophy* **17**, 247–279.

Nebel J. 2020. Rank-weighted utilitarianism and the veil of ignorance. *Ethics* **131**, 87–106.

Rawls J. 1996. *Political Liberalism*. New York: Columbia University Press.

Rawls J. 1999a. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rawls J. 1999b. *Collected Papers*, ed. S. Freeman. Cambridge, MA: Harvard University Press.

Risse M. 2002. Harsanyi's 'Utilitarian Theorem' and utilitarianism. *Nôus* **36**, 550–577.

Sen A. 1986. Social choice theory. In *Handbook of Mathematical Economics, volume 3*, ed. K.J. Arrow and M.D. Intriligator, 1073–1181. Amsterdam: North-Holland Publishing.

Thoma J. 2023. Taking risks on behalf of another. *Philosophy Compass* **18**, 1–13.

Thoma J. and J. Weinberg 2017. Risk writ large. *Philosophical Studies* **174**, 2369–2384.

van Fossen S. 2024. Can relative prioritarianism accommodate the shift? *Ethics* **134**, 525–538.

Weithman P. 2017. Autonomy and disagreement about justice in *Political Liberalism*. *Ethics* **128**, 95–122.

Weymark J. 1991. A reconsideration of the Harsanyi-Sen debate on utilitarianism. In *Interpersonal Comparisons of Well-Being*, ed. J. Elster and J. Roemer, 255–320. Cambridge: Cambridge University Press.

Weymark J. 1995. John Harsanyi's contributions to social choice and welfare economics. *Social Choice and Welfare* **12**, 313–318.

[9] Or it may not be (see Thoma 2023).

https://doi.org/10.1017/S0266267124000452 Published online by Cambridge University Press

Paul Weithman is the Glynn Family Honors Professor of Philosophy at Notre Dame, the university from which he received his BA in 1981. He earned his PhD at Harvard, where he wrote his dissertation under John Rawls and Judith Shklar. He joined the Notre Dame faculty in 1991. Professor Weithman works primarily in political philosophy but has also worked in moral philosophy, religious ethics, medieval political theory and the philosophy of education. He chaired the Philosophy Department between 2001 and 2007 and currently directs the interdisciplinary minor in Philosophy, Politics and Economics. He serves on the editorial boards of the *Journal of Religious Ethics*, *The Review of Politics* and *Politics, Philosophy and Economics*. He is an honorary member of the Brazilian Society for Legal Philosophy. His book *Why Political Liberalism?* won the David and Elaine Spitz prize as the best book in liberal and democratic theory published in 2010. URL: www.paulweithman.com

Cite this article: Weithman P. Avoiding risks behind the veil of ignorance. *Economics and Philosophy*. https://doi.org/10.1017/S0266267124000452