Original Position Arguments: An Axiomatic Characterization

October 18, 2022

Abstract

We study original position arguments in the context of social choice under ignorance. First, we present a general formal framework for such arguments. Next, we provide an axiomatic characterization of social choice rules that can be grounded by them. We illustrate this characterization in terms of various well-known social choice rules, some of which do and some of which do not satisfy the axioms in question. Depending on the perspective one takes, our results can be used to argue against certain rules, or conversely, against theories of procedural fairness that are built on the Rawlsian notion of original position arguments.

Keywords: original position arguments, social choice under ignorance, axiomatization.

1 Introduction

When a social planner decides between different policies, we expect her to employ a decision procedure or rule that is fair to the individuals whose welfare is affected by the decision.[1] Naturally, there are many competing views on what it means to say a procedure is fair. One prominent view takes an *ex ante* perspective and argues that "a procedure is fair if all parties would have agreed to the procedure had they been able to contract for it in advance of ("ex ante") their dispute" [4, p. 491]. The central question is then: why should such hypothetical consent be enough to justify imposing a procedure on someone who objects to it? Bone [4] distinguishes between two forms of contractarian theories that provide an answer: *egoistic contractarianism* and *ideal contractarianism.*

Egoistic contractarianism states that a person should comply with a procedure because if she were perfectly rational and well-informed she would have agreed to the procedure. By contrast, ideal contractarianism states that a person should comply because if she were put in an idealized choice situation, then it would be in her self-interest to accept the procedure. A prominent example of

[1] While the terms 'rule' and 'procedure' have different meanings, we treat them as interchangeable in this paper.

this last approach is Rawls's original position argument in *A Theory of Justice* [22].

The original position is a hypothetical situation where "no one knows his place in society, his class position or social status [...]" [22, p. 11]. Once we are placed in such a situation, the problem of ascertaining the fairness of a procedure is reduced to one of rational choice:

> Understood in this way the question of justification is settled by working out a problem of deliberation: we have to ascertain which principles it would be rational to adopt given the contractual situation [viz. the original position]. This connects the theory of justice with the theory of rational choice. [22, p. 16]

Importantly, even if we buy into ideal contractarianism, it remains an open question what principles we should adopt in the original position. Rawls, for one, argues that individuals in the original position are fully ignorant and have no reasonable basis to assign probabilities to the various possible outcomes [22, pp. 134-135]. He argues, moreover, that rational individuals would reason according to the maximin principle, which compares choices by looking only at their worst possible outcomes. This, it is argued, corresponds with the recommendations of the *difference principle*, which states that we should "arrange social and economic inequalities in such a way that they are to the benefit of the least advantaged" [22, p. 20].

The difference principle has been criticized by Sen [25] on the grounds that it violates the strong pareto principle.[2] Some have proposed a lexical variant of the difference principle, which says that one should first maximize the welfare of the worst-off individuals and then, in case of equal welfare, maximize the welfare of the second worst-off individuals, and so on. Both Parfit [20] and van Parijs [28] have claimed that Rawls's original position argument is better understood as supporting such a lexical difference principle.[3] In contrast to these proposals, Harsanyi [11, 12] argues that when faced with complete ignorance, we should assign every outcome an equal probability and maximize expected utility. This corresponds to the *principle of average utility*, which favours the options that lead to the highest average utility of the members of society.

The dispute between Rawls and Harsanyi has spawned a rich literature on social choice and decision-making under uncertainty [17, 19, 24, 10, 8, 5, 26, 6]. Emerging from this is the view that, even if they are not able to single out a unique principle of justice, original position arguments still serve as a useful tool in sorting out our intuitions regarding procedural fairness and its relation to social choice [14]. What is lacking, however, is an exact and general characterization of exactly which conceptions of justice or social choice can be

[2] On the strong pareto principle, a welfare distribution is optimal if and only if there is no other distribution such that no one is worse off and at least one individual is strictly better off under that other distribution.

[3] See [7] for a critical survey of lexical variants of the difference principle from the viewpoint of original position arguments.

grounded in an original position argument. The present paper contributes to filling this gap.

This paper Within the decision theoretic literature, choice under uncertainty is typically split into two types: *choice under ignorance* and *choice under risk*. The latter refers to cases where we know the probabilities of each possible state, whereas the former refers to cases where such information is absent [23, 21]. In this paper, we focus on social choice rules for decision-making under ignorance. We ask under what conditions original position arguments such as Rawls's and Harsanyi's can be successful. That is, instead of arguing for or against particular individual and social choice rules, we provide an axiomatic characterization of the general class of social choice rules that can be grounded in an original position argument.

We start by introducing standard models of (social) choice under ignorance and define the general classes of individual and social choice rules (Section 2). In Section 3, we present a general format for evaluating original position arguments, give examples of such arguments, and introduce the two axioms that make way for our central characterization theorem, which is then established in Section 4. This characterization result roughly says that a social choice rule can be grounded in an original position argument if and only if according to this social choice rule it does not matter which individual gets what and under what circumstances. We end with a discussion of our results and what they imply for original position arguments (Section 5).

Related work Maskin [15] gives a general, axiomatic characterization of individual choice rules under ignorance. As he indicates, these axiomatizations are strongly linked to results in social choice theory, but Maskin does not consider the issue of social choice under ignorance *per se*, let alone the Rawlsian notion of an original position.

Strasnick [27] also approaches original position arguments from an axiomatic angle. He argues that the concept of an original position entails a specific requirement on (a social planner's) priorities over individual preferences that, combined with various plausible principles of social choice, results in a ranking of distributions that agrees with the difference principle. It is an open question whether and how these insights can be generalized to deal with choice under uncertainty.

A general format for evaluating original position arguments within the context of choice under ignorance was first proposed in [7]. We use the same format and make the underlying assumptions about individual and social choice fully precise in this paper. Moreover, whereas [7] focuses on the lexical difference principle and original position arguments for it, our focus here is on axiomatizing a general class of social choice rules, viz. those that can be grounded in an original position argument.

2 Choice scenarios and choice rules

We start by introducing models of social choice under ignorance (Section 2.1). Once this is in place, we specify plausible individual and social choice rules and delineate a general class of individual and social choice rules (Section 2.2).

2.1 Choice scenarios

The models we use are introduced in [7]. These models are obtained by combining ingredients from the study of welfare distributions (cf. [25]) and from decision-making under ignorance (cf. [23, 21]).

Definition 1. *A* choice scenario *is a tuple* $\mathfrak{C} = \langle N, A, S, d\rangle$*, where* N *is a non-empty finite set of* individuals*,* A *a non-empty finite set of* alternatives*,* S *a non-empty finite set of* states*, and* $d : N \times A \times S \to \mathbb{R}$ *a* welfare distribution function*.*

In a given choice scenario, the set S represents the ignorance of the decision-maker, i.e. S is the set of states the decision-maker considers possible. The members of $A \times S$ are called the (possible) *outcomes* of the scenario. For each of these outcomes (a, s), the distribution function d determines the welfare of each individual $i \in N$. Depending on the application, one may require an ordinal or cardinal welfare scale and expect these values to be interpersonally comparable. In all our examples we assume interpersonal comparability and depending on the example a cardinal scale is also assumed. However, none of our technical results depend on any scale or interpretation of welfare. We work with finite models to keep our examples simple, but our characterization result does not depend on this assumption.

Figure 1 represents a simple choice scenario with two individuals 1 and 2, three alternatives, and two states. Here, the couples (n, m) represent the distribution function, where $n = d(1, a, s)$ and $m = d(2, a, s)$. For example, at outcome (b, s_2) individual 1's welfare is 1 whereas individual 2's welfare is 3, such that individual 2 is considered better off than individual 1 at that outcome. Throughout this article, we use this choice scenario (and variations of it) as our running example.

$\mathfrak{C}_1$	s_1	s_2
a	$(1, 1)$	$(2, 2)$
b	$(1, 2)$	$(1, 3)$
c	$(6, 0)$	$(0, 2)$

Figure 1: Choice scenario $\mathfrak{C}_1$, with two individuals, two states, and three alternatives.

2.2 Choice Rules

We consider two types of choice rules, viz. *social choice rules* and *individual choice rules*. Given a choice scenario $\mathfrak{C} = \langle N, A, S, d\rangle$, a social choice rule $\mathbf{S}$ determines a set of admissible alternatives $\mathbf{S}(\mathfrak{C}) \subseteq A$, whereas an individual choice rule $\mathbf{R}$ determines a set of admissible alternatives $\mathbf{R}(\mathfrak{C}, i) \subseteq A$ for each individual $i \in N$. In what follows, we provide examples of individual (Section 2.2.1) and social (Section 2.2.3) choice rules and give exact definitions of these concepts (Sections 2.2.2 and 2.2.4).

2.2.1 Examples of individual choice rules

Let us briefly recall two well-known individual choice rules to set the stage for later discussions. First, the *maximin* rule tells us to choose any alternative that maximizes the value of the worst possible outcome.[4] More precisely, where $\mathsf{min}(X)$ denotes the $\leq$-minimal element of a set X of real numbers, we have:

Definition 2 (Maximin admissibility)**.** *Where* $\mathfrak{C} = \langle N, A, S, d\rangle$ *is a choice scenario,* $i \in N$*, and* $a \in A$*:* $a \in \mathbf{R}^{\mathsf{m}}(\mathfrak{C}, i)$ *iff for all* $b \in A : \mathsf{min}\{d(i, a, s) \mid s \in S\} \geq \mathsf{min}\{d(i, b, s) \mid s \in S\}$.

For example, in Figure 1, alternatives a and b are maximin admissible for individual 1 whereas for individual 2 only alternative b is admissible.[5]

Second, the *expected utility* rule tells us to choose any alternative that maximizes expected utility. However, recall that we do not assume that individuals have expectations about the relative likelihood of states. In order to perform expected utility calculations, we rely on the *principle of insufficient reason* or *principle of indifference* (cf. [13]), which states that in the absence of relevant evidence, individuals should assume that every state is equally likely.

Notation 1. *Let* $\mathfrak{C} = \langle N, A, S, d\rangle$ *be a choice scenario. Where* $a \in A$ *and* $i \in N$*, we write* $\mathsf{eu}_i(a)$ *to denote the expected utility of* a *for* i*, i.e.*

$$\mathsf{eu}_i(a) = \sum_{s \in S} \frac{d(i, a, s)}{|S|}$$

.

Definition 3 (Expected utility admissibility)**.** *Where* $\mathfrak{C} = \langle N, A, S, d\rangle$ *is a choice scenario,* $i \in N$*, and* $a \in A$*:* $a \in \mathbf{R}^{\mathsf{eu}}(\mathfrak{C}, i)$ *iff for all* $b \in A : \mathsf{eu}_i(a) \geq \mathsf{eu}_i(b)$.

In our running example (Figure 1), we have $\mathsf{eu}_1(a) = 1.5$, $\mathsf{eu}_1(b) = 1$, and $\mathsf{eu}_1(c) = 3$. Hence, the ranking for individual 1 induced by expected utility is

[4] See [15] for an axiomatic characterization of the maximin rule within the context of choice under ignorance.

[5] Maximin is often described as a conservative rule as it only takes into account the worst outcomes [21]. To remedy this, a number of more sophisticated rules have been proposed, such as the leximin and optimism-pessimism rule. However, we do not consider these rules in this paper.

$c \succ a \succ b$. For individual 2, we have $b \succ a \succ c$ since $\mathsf{eu}_2(a) = 1.5$, $\mathsf{eu}_2(b) = 2.5$, and $\mathsf{eu}_2(c) = 1$. In conclusion, only alternative c is expected utility admissible for individual 1, whereas only alternative b is expected utility admissible for individual 2.

2.2.2 An exact characterization of individual choice rules

In general, we define individual choice rules as *pointed* choice rules that satisfy certain axioms. By a pointed choice rule we mean a choice rule that determines a set of admissable alternatives relative to a specific "point", i.e. an individual.

Definition 4 (Pointed choice rule)**.** $\mathbf{R}$ *is a* pointed choice rule *iff for each choice scenario* $\mathfrak{C} = \langle N, A, S, d \rangle$ *and individual* $i \in N$, $\mathbf{R}(\mathfrak{C}, i) \subseteq A$.

An individual choice rule is a pointed choice rule that satisfies *Individualism* and *Column Symmetry*. First, *Individualism* (I) requires that for each individual, its set of admissible alternatives does not depend on the payoffs of other individuals in the same scenario. This implies that whatever is admissible for one individual does not change when we change the payoffs of other individuals.

Definition 5 (i-equivalence)**.** *Let* $\mathfrak{C} = \langle N, A, S, d \rangle$ *and* $\mathfrak{C}' = \langle N, A, S, d' \rangle$ *be choice scenarios and let* $i \in N$*. Scenarios* $\mathfrak{C}$ *and* $\mathfrak{C}'$ *are* i*-equivalent iff for all* $a \in A$, $s \in S : d'(i, a, s) = d(i, a, s)$.

In Figure 2, scenarios $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are 1-equivalent because individual 1 receives exactly the same payoffs in $\mathfrak{C}_1$ as in $\mathfrak{C}_2$. By contrast, individual 2 receives different payoffs in both scenarios. Individualism requires that if a is admissible for individual 1 in $\mathfrak{C}_1$, then a should be admissible for individual 1 in $\mathfrak{C}_2$ as well.

$\mathfrak{C}_1$	s_1	s_2
a	$(1,1)$	$(2,2)$
b	$(1,2)$	$(1,3)$
c	$(6,0)$	$(0,2)$

$\mathfrak{C}_2$	s_1	s_2
a	$(1,0)$	$(2,0)$
b	$(1,0)$	$(1,0)$
c	$(6,0)$	$(0,0)$

Figure 2: Scenario $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are 1-equivalent but not 2-equivalent.

Individualism If $\mathfrak{C}$ and $\mathfrak{C}'$ are i-equivalent then $\mathbf{R}(\mathfrak{C}', i) = \mathbf{R}(\mathfrak{C}, i)$.

Second, *Column Symmetry* (CS) requires that a choice rule is not sensitive to the way states are labelled. This principle is one half of Milnor's [16] Symmetry Condition, which states that the labelling of states and alternatives should be irrelevant to decision criteria.[6] It is considered standard when dealing with choice under ignorance (cf. [21]). The following definition determines when we can relabel the states from different scenarios to obtain one from the other.

[6] Milnor's Symmetry Condition is also known as Arrow and Hurwicz's property B [1].

Definition 6 (S-label variants)**.** *Let* $\mathfrak{C} = \langle N, A, S, d\rangle$ *and* $\mathfrak{C}' = \langle N, A, S', d'\rangle$ *be choice scenarios. Scenarios* $\mathfrak{C}$ *and* $\mathfrak{C}'$ *are* S-label variants *iff there exists a bijection* $\sigma : S \to S'$ *such that for all* $i \in N$, $a \in A$ *and* $s \in S : d'(i, a, \sigma(s)) = d(i, a, s)$.

In Figure 3, we see two choice scenarios that are S-label variants. For both individuals, the payoffs they receive at s_1 in $\mathfrak{C}_1$ are identical to the payoffs they receive in s'_1 in $\mathfrak{C}_2$, and the same holds for s_2 and s'_2. Column Symmetry requires that if a is admissible for individual 1 in $\mathfrak{C}_1$, then a should also be admissible for individual 1 in $\mathfrak{C}_2$.

$\mathfrak{C}_1$	s_1	s_2
a	$(1,1)$	$(2,2)$
b	$(1,2)$	$(1,3)$
c	$(6,0)$	$(0,2)$

$\mathfrak{C}_2$	s'_1	s'_2
a	$(1,1)$	$(2,2)$
b	$(1,2)$	$(1,3)$
c	$(6,0)$	$(0,2)$

Figure 3: Scenario $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are S-label variants.

Column Symmetry If $\mathfrak{C}$ and $\mathfrak{C}'$ are S-label variants then for all $i \in N : \mathbf{R}(\mathfrak{C}, i) = \mathbf{R}(\mathfrak{C}', i)$.

We submit that both Individualism and Column Symmetry are the weakest requirements one can impose on pointed choice rules without compromising the concept of an original position argument. First, Column Symmetry is both common and plausible. More importantly, as shown in Appendix A, without this axiom, the notion of an original position argument as we characterize it is trivialized (Theorem 4). Second, Individualism is integral to the concept of an original position argument.[7] Rawls writes: "The essential idea is that we want to account for the social values, for the intrinsic good of institutional, community, and associative activities, by a conception of justice that in its theoretical basis is individualistic." [22, p. 233]. Hence, the attractiveness of original position arguments relies on the fact that some social choice rules can be reduced to the rational decision-making of self-interested persons (i.e. they are individualistic) under hypothetical circumstances considered fair.

In what follows, we take individual choice rules to be pointed choice rules that satisfy *Individualism* and *Column Symmetry*. Clearly, both maximin and the expected utility rule satisfy Individualism since both determine a set of admissible alternatives for each individual $i \in N$ and only take into account the payoffs of that individual. Likewise, both rules satisfy Column Symmetry for obvious reasons. Hence, maximin and the expected utility rule are individual choice rules in the sense just defined.

[7] It is possible to give up Individualism without trivializing the concept of an original position argument. In Appendix A, we state and prove the characterization result that arises in such a setting (Corollary 3).

2.2.3 Examples of Social Choice Rules

In this section, we introduce four distinct social choice rules. These social choice rules will be used as examples throughout the article.

Difference Principle The difference principle states that we should "arrange social and economic inequalities in such a way that they are to the benefit of the least advantaged" [22, p. 20]. Conceived as a social choice rule, the difference principle tells us to choose any alternative that maximizes the prospects of the least well-off. However, once we are dealing with a context of choice under ignorance, it is ambiguous what exactly "least well-off" means, and hence one may specify this principle in conceptually distinct ways (see [7] for an overview of these approaches). Here, we will focus on one approach, viz. what is called the basic approach in [7]. On the basic approach, we maximize welfare, ignoring the distinction between different states and different individuals.

Definition 7 (Difference admissibility)**.** *Where* $\mathfrak{C} = \langle N, A, S, d\rangle$ *is a choice scenario and* $a \in A$*:* $a \in \mathbf{S}^{\mathsf{d}}(\mathfrak{C})$ *iff for all* $b \in A : \mathsf{min}\{d(i,a,s) \mid i \in N \,\&\, s \in S\} \geq \mathsf{min}\{d(i,b,s) \mid i \in N \,\&\, s \in S\}$*.*

For example, in the scenario in Figure 1 on page 4, we have $\mathsf{min}\{d(i,a,s) \mid i \in N \,\&\, s \in S\} = 1$, $\mathsf{min}\{d(i,b,s) \mid i \in N \,\&\, s \in S\} = 1$, and $\mathsf{min}\{d(i,c,s) \mid i \in N \,\&\, s \in S\} = 0$. Hence, both a and b are difference admissible, but c is not.

Average Expected Utility The second social choice rule that we discuss is the *average expected utility* rule. It tells us to choose any alternative that maximizes the average expected utility of all individuals.

Notation 2. *Let* $\mathfrak{C} = \langle N, A, S, d\rangle$ *be a choice scenario. Where* $a \in A$*, we write* $\mathsf{aeu}(a)$ *to denote the average expected utility of* a*, i.e.*

$$\mathsf{aeu}(a) = \frac{\sum_{i \in N} \mathsf{eu}_i(a)}{|N|}$$

.

Definition 8 (Average expected utility admissibility)**.** *Where* $\mathfrak{C} = \langle N, A, S, d\rangle$ *is a choice scenario and* $a \in A$*:* $a \in \mathbf{S}^{\mathsf{aeu}}(\mathfrak{C})$ *iff for all* $b \in A : \mathsf{aeu}(a) \geq \mathsf{aeu}(b)$*.*

Applying the average expected utility rule to our running example (cf. Figure 1), we have $\mathsf{aeu}(a) = 1.5$, $\mathsf{aeu}(b) = 1.75$, and $\mathsf{aeu}(c) = 2$, and hence the ranking induced by average expected utility is $c \succ b \succ a$. Notice that the expected utility for individual 1 under c is very high ($\mathsf{eu}_1(c) = 3$), whereas for individual 2 it is relatively low ($\mathsf{eu}_2(c) = 1$). Still, the rule picks c because the average social utility is skewed upwards by the great prospects of individual 1, even though it is the worst alternative for individual 2. For this reason, the average expected utility rule is sometimes criticized on the grounds that it allows for the "sacrificing" of those who are less well-off if doing so would be offset by a sufficient benefit to others. In order to remedy this, one should give greater weight to the expectations of those who are less well-off. More generally, one may consider ways of combining the difference principle and expected utility. In what follows, we consider two such combinations originally introduced by Mongin and Pivato in [18].[8]

Difference Expected Utility According to this social choice rule we maximize the social value of alternatives, where the social value of an alternative is the expected utility of the individual who has the worst prospects under that alternative.

Definition 9 (Difference expected utility admissibility)**.** *Where* $\mathfrak{C} = \langle N, A, S, d\rangle$ *is a choice scenario and* $a \in A$*:* $a \in \mathbf{S}^{\mathsf{deu}}(\mathfrak{C})$ *iff for all* $b \in A : \min\{\mathsf{eu}_i(a) \mid i \in N\} \geq \min\{\mathsf{eu}_i(b) \mid i \in N\}$*.*

In our running example (cf. Figure 1), we have $\min\{\mathsf{eu}_i(a) \mid i \in N\} = 1.5$, $\min\{\mathsf{eu}_i(b) \mid i \in N\} = 1$, and $\min\{\mathsf{eu}_i(c) \mid i \in N\} = 1$. Hence, the difference expected utility ranking is $a \succ b = c$.

Maximin Expected Utility On this social choice rule, to determine the social value of an alternative, we obtain for each outcome the value of the worst-off person at that outcome and then apply the expected utility rule to these values.

Notation 3. *Let* $\mathfrak{C} = \langle N, A, S, d\rangle$ *be a choice scenario. Where* $a \in A$*, we write* $\mathsf{meu}(a)$ *to denote the maximin expected utility of alternative* a*, i.e.*

$$\mathsf{meu}(a) = \sum_{s \in S} \frac{\min\{d(i, a, s) \mid i \in N\}}{|S|}$$

.

Definition 10 (Maximin expected utility admissibility)**.** *Where* $\mathfrak{C} = \langle N, A, S, d\rangle$ *is a choice scenario and* $a \in A$*,* $a \in \mathbf{S}^{\mathsf{mau}}(\mathfrak{C})$ *iff for all* $b \in A : \mathsf{meu}(a) \geq \mathsf{meu}(b)$*.*

Looking at our running example once more, we have $\mathsf{meu}(a) = 1.5$, $\mathsf{meu}(b) = 1$, and $\mathsf{meu}(c) = 0$. Hence, the induced ranking is $a \succ b \succ c$.

2.2.4 An exact characterization of social choice rules

We introduce the general class of social choice rules. Recall that these are conceived as rules that map every choice scenario $\mathfrak{C} = \langle N, A, S, d\rangle$ to a subset $\mathbf{S}(\mathfrak{C}) \subseteq A$ of socially admissible alternatives. In addition to this formal requirement, we require that social choice rules satisfy *State Multiplication Indifference*. This axiom says that if we duplicate every state of a given scenario a fixed number of times, then the admissibility of an alternative should be preserved.

[8] Mongin and Pivato [18] write: "probabilities can enter the maximin rule in accordance with two different methods. Either an expected value is first taken for each individual and maximin is then applied, which is the ex ante method, or maximin is first applied in each state and the expected value is then taken, which is the ex post method.". The difference expected utility rule we mention corresponds to what they call the *ex ante* approach to maximin, and the maximin expected utility rule to the *ex post* approach.

	$\mathbf{S}^{\mathsf{d}}$	$\mathbf{S}^{\mathsf{aeu}}$	$\mathbf{S}^{\mathsf{deu}}$	$\mathbf{S}^{\mathsf{meu}}$
a	1	1.5	1.5	1.5
b	1	1.75	1	1
c	0	2	1	0
	$a = b \succ c$	$c \succ b \succ a$	$a \succ b = c$	$a \succ b \succ c$

Table 1: The choice rules applied to the running example (cf. Figure 1).

Definition 11 (State Multiplication variants)**.** *Let $\mathfrak{C} = \langle N, A, S, d\rangle$ be a choice scenario and let Δ be a non-empty finite set. Scenarios $\mathfrak{C}$ and $\mathfrak{C}^{\Delta} = \langle N, A, S^{\Delta}, d^{\Delta}\rangle$ are State Multiplication variants iff*

- $S^{\Delta} = S \times \Delta$
- *for all $i \in N$, $a \in A$, $(s, \delta) \in S^{\Delta}$: $d^{\Delta}(i, a, (s, \delta)) = d(i, a, s)$.*

State Multiplication Indifference For all choice scenarios $\mathfrak{C}$ and all non-empty and finite Δ: $\mathbf{S}(\mathfrak{C}) = \mathbf{S}(\mathfrak{C}^{\Delta})$.

All the social choice rules introduced above satisfy State Multiplication Indifference. For example, for the expected average utility rule, multiplying the entire set of states does not change the expected utility of any individual. Hence, since the expected utility of each individual does not change after multiplication, the expected average utility also does not change. State Multiplication Indifference is weaker than the property known as Column Duplication[9] [3, 16]. The latter allows for adding copies of just some states, whereas State Multiplication only allows one to copy all states a given number of times.

Definition 12 (Social choice rule)**.** $\mathbf{S}$ *is a* social choice rule *iff for each choice scenario $\mathfrak{C} = \langle N, A, S, d\rangle$, $\mathbf{S}(\mathfrak{C}) \subseteq A$ and $\mathbf{S}$ satisfies* State Multiplication Indifference.

Our notion of social choice rules is very liberal: it assumes almost no properties. This is as intended. In what follows we ask, within this very broad class, which social choice rules can be grounded in an original position argument.

3 Original Position Arguments

What does it mean that a given individual choice rule can be used to ground a social choice rule, using an original position argument? We start by giving a definition of original position arguments within our format (Section 3.1). Next, we give examples of such arguments for concrete social choice rules (Section 3.2). Finally, we introduce the axioms that characterize the class of all social choice rules that can be grounded in an original position argument (Section 3.3).

[9]Column Duplication is also known as State-Individuation Invariance (cf. [9]) or Independence of Duplicate States (cf. [2]).

3.1 A definition of original position arguments

We present the general format for evaluating original position arguments introduced in [7]. The starting point is the view that the original position does not correspond to a particular scenario: instead, for each particular choice scenario, we can construct a corresponding choice scenario which has the characteristics of an original position. The latter is called the *original position transformation* of the original choice scenario.

Definition 13 (OP-transformation)**.** *Let* $\mathfrak{C} = \langle N, A, S, d \rangle$ *be a choice scenario. Let* Π *be the set of all bijective functions* $\pi : N \to N$*. The* original position transformation *of* $\mathfrak{C}$ *is the choice scenario* $\mathfrak{C}^* = \langle N, A, S^*, d^* \rangle$*, where*

- $S^* = S \times \Pi$
- *for all* $i \in N$*,* $a \in A$*, and* $(s, \pi) \in S^*$*:* $d^*(i, a, (s, \pi)) = d(\pi(i), a, s)$

In other words, given some choice scenario $\mathfrak{C}$, we obtain its OP-transformation $\mathfrak{C}^*$ by combining the ignorance in the original model with ignorance about the individual's identities and the way these identities affect the level of welfare one receives. We illustrate this by means of our running example. Figure 4 displays (on the left-hand side) the choice scenario $\mathfrak{C}_1$, and (on the right-hand side) its OP-transformation. Here, $\pi_=$ is the identity relation, and $\pi_{\neq}$ swaps the two individuals, i.e. $\pi_=(1) = 1$, $\pi_=(2) = 2$, $\pi_{\neq}(1) = 2$, and $\pi_{\neq}(2) = 1$. If we apply the maximin rule to the OP-transformation, we find that both a and b are admissible for individual 1, while c is not.

$\mathfrak{C}_1$	s_1	s_2
a	$(1,1)$	$(2,2)$
b	$(1,2)$	$(1,3)$
c	$(6,0)$	$(0,2)$

$\mathfrak{C}^*_1$	$(s_1, \pi_=)$	$(s_2, \pi_=)$	$(s_1, \pi_{\neq})$	$(s_2, \pi_{\neq})$
a	$(1,1)$	$(2,2)$	$(1,1)$	$(2,2)$
b	$(1,2)$	$(1,3)$	$(2,1)$	$(3,1)$
c	$(6,0)$	$(0,2)$	$(0,6)$	$(2,0)$

Figure 4: A choice scenario ($\mathfrak{C}_1$) and its OP-transformation ($\mathfrak{C}^*_1$).

With this in place, we can give an exact definition of what it means for a social choice rule to be grounded in an original position argument.

Definition 14 (Original position derivation)**.** *Let* $\mathbf{S}$ *be a social choice rule, and let* $\mathbf{R}$ *be an individual choice rule.* $\mathbf{S}$ *can be original position derived from* $\mathbf{R}$ *iff for all scenarios* $\mathfrak{C} = \langle N, A, S, d \rangle$ *and all* $i \in N : \mathbf{S}(\mathfrak{C}) = \mathbf{R}(\mathfrak{C}^*, i)$*.*
The social choice rule $\mathbf{S}$ *can be grounded in an original position argument iff there is an individual choice rule* $\mathbf{R}$ *such that* $\mathbf{S}$ *is original position derived from* $\mathbf{R}$.

3.2 Two examples of original position arguments

Recall that both Rawls and Harsanyi claimed that their favoured social choice rules can be supported by an original position argument. We will show that we can verify counterparts of these claims, in the context of choice under ignorance. In particular, we will show that the difference principle can be OP-derived from the maximin rule (Proposition 1) and that the principle of average expected utility can be OP-derived from the expected utility rule (Proposition 2). We also provide an example of a social choice rule that cannot be grounded in an original position argument: the difference expected utility rule (Proposition 3).

Before we get to the examples, it will be helpful to introduce some extra notation. In particular, we work with multisets, i.e. sets that can contain multiple instances of the same member. To distinguish a multiset from a regular set, we use rectangular brackets [,] instead of {, }.

We start by observing that there is a specific relation between the payoffs of all individuals in a choice scenario and the payoffs of a fixed individual in its original position transformation:

Fact 1. *For all choice scenarios* $\mathfrak{C} = \langle N, A, S, d\rangle$, $i \in N$ *and* $a \in A$*:*

$$[d(j, a, s) \mid j \in N \,\&\, s \in S] = [d(i, a, s) \mid s \in S^*]$$

Fact 1 says that, given an alternative, the multiset of possible payoffs *any* individual can receive in a choice scenario is identical to the multiset of possible payoffs a fixed individual can receive in its OP-transformation. This holds because the OP-transformation of any scenario is constructed precisely so that each individual in the original position considers the possibility of receiving the payoffs of each individual in the pre-transformed scenario.

Proposition 1. *The difference principle can be OP-derived from the maximin rule.*

Proof. We show that for all choice scenarios $\mathfrak{C} = \langle N, A, S, d\rangle$ and all $i \in N$: $\mathbf{R}^{\mathsf{m}}(\mathfrak{C}^*, i) = \mathbf{S}^{\mathsf{d}}(\mathfrak{C})$. Fix a choice scenario $\mathfrak{C} = \langle N, A, S, d\rangle$, let $i \in N$ and $a \in A$ be arbitrary. We have $a \notin \mathbf{R}^{\mathsf{m}}(\mathfrak{C}^*, i)$ iff [By Definition 2] there is a $b \in A$ such that $\mathsf{min}\{d(i, b, s) \mid s \in S^*\} > \mathsf{min}\{d(i, a, s) \mid s \in S^*\}$ iff [By Fact 1] there is a $b \in A$ such that $\mathsf{min}\{d(j, b, s) \mid j \in N \,\&\, s \in S\} > \mathsf{min}\{d(j, a, s) \mid j \in N \,\&\, s \in S\}$ iff [By Definition 7] $a \notin \mathbf{S}^{\mathsf{d}}(\mathfrak{C})$. □

Proposition 2. *The average expected utility principle can be OP-derived from the expected utility rule.*

Proof. We show that for all choice scenarios $\mathfrak{C} = \langle N, A, S, d\rangle$ and all $i \in N$: $\mathbf{R}^{\mathsf{eu}}(\mathfrak{C}^*, i) = \mathbf{S}^{\mathsf{aeu}}(\mathfrak{C})$. Fix a choice scenario $\mathfrak{C} = \langle N, A, S, d\rangle$, let $i \in N$ and $a \in A$ be arbitrary. We have $a \notin \mathbf{R}^{\mathsf{eu}}(\mathfrak{C}^*, i)$ iff [By Definition 3] there is $b \in A$ such that $\mathsf{eu}_i(b) > \mathsf{eu}_i(a)$ iff [Notation 1 and simplifying the expression] $\sum[d(i, b, s) \mid s \in S^*] > \sum[d(i, a, s) \mid s \in S^*]$ iff [By Fact 1] there is $b \in A$ such that $\sum[d(j, b, s) \mid j \in N \,\&\, s \in S] > \sum[d(j, a, s) \mid j \in N \,\&\, s \in S]$ iff [Notation 2] there is $b \in A$ such that $\mathsf{aeu}(b) > \mathsf{aeu}(a)$ iff [By Definition 8] $a \notin \mathbf{S}^{\mathsf{aeu}}(\mathfrak{C})$. □

Proposition 3. *The difference expected utility principle cannot be grounded in an original position argument.*

Proof. We show that there is no individual choice rule $\mathbf{R}$ such that for all choice scenarios $\mathfrak{C} = \langle N, A, S, d\rangle$ and all $i \in N$: $\mathbf{R}(\mathfrak{C}^*, i) = \mathbf{S}^{\mathsf{deu}}(\mathfrak{C})$. Consider the following scenarios:

$\mathfrak{C}_1$	s_1	s_2
a	$(0,1)$	$(0,1)$
b	$(0,0)$	$(0,0)$

$\mathfrak{C}_1^*$	$(s_1, \pi_=)$	$(s_2, \pi_=)$	$(s_2, \pi_{\neq})$	$(s_2, \pi_{\neq})$
a	$(0,1)$	$(0,1)$	$(1,0)$	$(1,0)$
b	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$

$\mathfrak{C}_2$	s_1	s_2
a	$(0,1)$	$(1,0)$
b	$(0,0)$	$(0,0)$

$\mathfrak{C}_2^*$	$(s_1, \pi_=)$	$(s_2, \pi_=)$	$(s_1, \pi_{\neq})$	$(s_2, \pi_{\neq})$
a	$(0,1)$	$(1,0)$	$(1,0)$	$(0,1)$
b	$(0,0)$	$(0,0)$	$(0,0)$	$(0,0)$

Compare scenarios $\mathfrak{C}_1$ and $\mathfrak{C}_2$. In scenario $\mathfrak{C}_1$, both alternative a and alternative b are admissible according to $\mathbf{S}^{\mathsf{deu}}$, since the expected utility of the worst-off person under both is 0. By contrast, in $\mathfrak{C}_2$ only alternative a is admissible since the expected utility of both individuals is 0.5, whereas the expected utility for each under alternative b is 0. Hence, $\mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_1) \neq \mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_2)$.

The proof now proceeds by reductio. Suppose that there is some individual choice rule $\mathbf{R}$ that grounds the social choice rule $\mathbf{S}^{\mathsf{deu}}$. Hence, $x \in \mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_1)$ iff $x \in \mathbf{R}(\mathfrak{C}_1^*)$. Similary for scenario $\mathfrak{C}_2$ we have $x \in \mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_2)$ iff $x \in \mathbf{R}(\mathfrak{C}_2^*)$. By Column Symmetry and Individualism, $x \in \mathbf{R}(\mathfrak{C}_1^*)$ iff $x \in \mathbf{R}(\mathfrak{C}_2^*)$. Following this chain of equivalences allows us to conclude that $x \in \mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_1)$ iff $x \in \mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_2)$, which contradicts our earlier observation that $\mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_1) \neq \mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_2)$.

□

So far, we have illustrated how one can show that a social choice rule can be grounded in an original position argument (Propositions 1 and 2) as well as how one can argue that it is impossible to do so (Proposition 3).[10] In doing so, we had to rely on the specific properties of the social choice rules in question. However, one may also ask whether there are general properties that make original position arguments tick. This would allow us to reduce the question of whether a given social choice rule can be grounded to the question of whether it satisfies these properties.

3.3 Indifference axioms

In this section, we introduce and discuss two axioms that we show to be characteristic of original position arguments. First, *Indifference to Intra-State Distribution of Payoffs to Persons* (IISD) states that social admissibility should not

[10] One can similarly show that the maximin expected utility rule cannot be grounded in an original position argument.

depend on how the payoffs within states are distributed across individuals. To put it plainly, given any particular outcome, it should not matter whether Bob gets a cake and Alice an apple or the other way around. Let us make this more precise.

Definition 15 (Π-variants). *Let* $\mathfrak{C} = \langle N, A, S, d\rangle$ *and* $\mathfrak{C}' = \langle N, A, S, d'\rangle$ *be choice scenarios.* $\mathfrak{C}$ *and* $\mathfrak{C}'$ *are* Π-variants *iff for all* $s \in S$ *there is a bijection* $\pi_s : N \to N$ *such that for all* $a \in A$, $i \in N$: $d'(\pi_s(i), a, s) = d(i, a, s)$.

Figure 5 below depicts two choice scenarios that are Π-variants of each other. Note that in scenario $\mathfrak{C}_2$ the payoffs of individuals 1 and 2 are switched at s_1 compared to s_1 in $\mathfrak{C}_1$, i.e. $\pi_{s_1}(i) = j, \pi_{s_1}(j) = i$ and $\pi_{s_2}(i) = i, \pi_{s_2}(j) = j$. In general, for different states s and s' one may have different permutations: $\pi_s \neq \pi_{s'}$.

$\mathfrak{C}_1$	s_1	s_2
a	$(1,1)$	$(2,2)$
b	$(1,2)$	$(1,3)$
c	$(6,0)$	$(0,2)$

$\mathfrak{C}_2$	s_1	s_2
a	$(1,1)$	$(2,2)$
b	$(2,1)$	$(1,3)$
c	$(0,6)$	$(0,2)$

Figure 5: Two Π variants.

Indifference to Intra-State Distribution of Payoffs to Persons If $\mathfrak{C}$ and $\mathfrak{C}'$ are Π-variants then $\mathbf{S}(\mathfrak{C}) = \mathbf{S}(\mathfrak{C}')$.

One way to interpret the IISD axiom is to view it as securing what might be called *ex post anonymity*; i.e. once a particular outcome is fixed, the labels of individuals should not matter. An example of a social choice rule that satisfies IISD is the maximin expected utility rule. The maximin expected utility rule is only sensitive to the utility values within states, and any permutation of the individual's payoffs at those states does not affect the utility values at those state. By contrast, the difference expected utility rule does not satisfy IISD. For example, in Figure 5, the difference expected utility rule considers alternative a admissible in scenario $\mathfrak{C}_1$, whereas both alternative a and alternative b are considered admissible in scenario $\mathfrak{C}_2$.

Our second axiom, *Indifference to Intra-Person Distribution of Payoffs to States* (IIPD), says that social admissibility should not depend on how each individual's payoffs are distributed across states. For example, it should not make a difference whether Bob gets a cake in state s and an apple in state s' but only that Bob either gets an apple or a cake.

Definition 16 (Σ-variants). *Let* $\mathfrak{C} = \langle N, A, S, d\rangle$ *and* $\mathfrak{C}' = \langle N, A, S, d'\rangle$ *be choice scenarios.* $\mathfrak{C}$ *and* $\mathfrak{C}'$ *are* Σ-variants *iff for all* $i \in N$ *there is a bijection* $\sigma_i : S \to S$ *such that for all* $a \in A$: $d'(i, a, \sigma_i(s)) = d(i, a, s)$.

Figure 6 depicts two Σ-variants. In scenario $\mathfrak{C}_2$, the payoffs of individual 1 are switched between s_1 and s_2 compared to scenario $\mathfrak{C}_1$, whereas the payoffs of individual 2 are untouched, i.e. $\sigma_1(s_1) = s_2, \sigma_1(s_2) = s_1$ and $\sigma_2(s_1) = s_1, \sigma_2(s_2) = s_2$.

$\mathfrak{C}_1$	s_1	s_2
a	$(1,1)$	$(2,2)$
b	$(1,2)$	$(1,3)$
c	$(6,0)$	$(0,2)$

$\mathfrak{C}_2$	s_1	s_2
a	$(2,1)$	$(1,2)$
b	$(1,2)$	$(1,3)$
c	$(0,0)$	$(6,2)$

Figure 6: Two Σ-variants.

Indifference to Intra-Person Distribution of Payoffs to States If $\mathfrak{C}$ and $\mathfrak{C}'$ are Σ-variants then $\mathbf{S}(\mathfrak{C}) = \mathbf{S}(\mathfrak{C}')$.

The difference expected utility rule satsifies IIPD, whereas the maximin expected utility rule does not. The difference expected utility rule satisfies IIPD because any permutation of an individuals' payoffs across states does not change the expected utility of that individual. Of course, this only holds because we are assuming that each state is equally likely and hence switching payoffs between equally likely states does not make a difference. To see that the maximin expected utility rule does not satisfy IIPD, consider Figure 6. In scenario $\mathfrak{C}_1$, only alternative a is maximin expected utility admissible, whereas in scenario $\mathfrak{C}_2$ all alternatives are admissible.

A little reflection on the definition of IIPD reveals that it implies Column Symmetry.[11] Suppose $\mathfrak{C}$ and $\mathfrak{C}'$ are S-label variants (cf. Definition 6). Hence, there is some $\sigma : S \to S$ such that for all $i \in N$, $a \in A$ and $s \in S : d'(i, a, \sigma(s)) = d(i, a, s)$. Given σ, we let $\sigma_i = \sigma$ for each $i \in N$. This gives us exactly what is needed to satisfy the definition of Σ-variants (Definition 16).

Fact 2. *If a social choice rule* $\mathbf{S}$ *satisfies IIPD, it satisfies Column Symmetry.*

We also note that IISD implies *Anonymity*, i.e. the axiom that tells us that the labels of individuals do not matter [25]. Since we need not rely on *Anonymity* for our results to go through, we will omit a detailed discussion of this principle. Taken together, IISD and IIPD say that it does not matter which individual gets what and under what circumstances. While these are rather strong conditions, both the basic difference rule and the average expected utility rule satisfy them.

4 Axiomatic Characterization

In this section, we prove that a social choice rule can be grounded in an original position argument if and only if it satisfies Indifference to Intra-Person Distribu-

[11] While technically we have not defined the Column Symmetry axiom for social choice rules, its definition is analogous to its definition for individual choice rules.

	$\mathbf{S}^{\mathsf{d}}$	$\mathbf{S}^{\mathsf{aeu}}$	$\mathbf{S}^{\mathsf{deu}}$	$\mathbf{S}^{\mathsf{mau}}$
State Multiplication Indifference	+	+	+	+
Indifference to Intra-State Distribution of Payoffs to Persons	+	+	−	+
Indifference to Intra-Person Distribution of Payoffs to States	+	+	+	−

Table 2: Overview of the properties of social choice rules.

tion of Payoffs to States and Indifference to Intra-State Distribution of Payoffs to Persons. In Section 4.1 we prove the implication from left to right and in Section 4.2 the implication from right to left.

4.1 Necessary conditions

We prove that if a social choice rule can be grounded in an original position argument, it satisfies IISD (Theorem 1) and IIPD (Theorem 2).

Our first lemma establishes that if two scenarios are Π-variants, then their OP-transformations agree on the set of admissible alternatives for every individual. In what follows, if f and g are functions, we use $g \circ f$ to denote the composition of f and g, i.e. $g \circ f(x) = g(f(x))$.

Lemma 1. *Let* $\mathbf{R}$ *be an individual choice rule. For all choice scenarios* $\mathfrak{C}_1 = \langle N, A, S, d_1\rangle$ *and* $\mathfrak{C}_2 = \langle N, A, S, d_2\rangle$*: if* $\mathfrak{C}_1$ *and* $\mathfrak{C}_2$ *are* Π*-variants, then for all* $i \in N : \mathbf{R}(\mathfrak{C}_1^*, i) = \mathbf{R}(\mathfrak{C}_2^*, i)$.

Proof. Let $\mathfrak{C}_1 = \langle N, A, S, d_1\rangle$ and $\mathfrak{C}_2 = \langle N, A, S, d_2\rangle$ be choice scenarios and let $\mathbf{R}$ be an individual choice rule. Suppose that $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are Π-variants. We show that $\mathfrak{C}_1^*$ and $\mathfrak{C}_2^*$ are S-label variants. By the supposition, for all $s \in S$ there is a bijection $\pi_s : N \to N$ such that for all $a \in A$ and $i \in N : d_2(\pi_s(i), a, s) = d_1(i, a, s)$ (Definition 15). We define $\sigma^* : S^* \to S^*$ as follows. Given some arbitrary state $(s, \pi) \in S^*$, let $\sigma^*(s, \pi) = (s, \pi_s \circ \pi)$. Let $i \in N$, $a \in A$, $(s, \pi) \in S^*$, and $m \in \mathbb{R}$ be arbitrary. We have:

$$
\begin{aligned}
d_1^*(i, a, (s, \pi)) = m & \\
\text{iff} \qquad & \text{Definition 13 (OP-transformation)} \\
d_1(\pi(i), a, s) = m & \\
\text{iff} \qquad & \text{Definition 15 (}\Pi\text{-variants)} \\
d_2(\pi_s \circ \pi(i), a, s) = m & \\
\text{iff} \qquad & \text{Definition 13 (OP-transformation)} \\
d_2^*(i, a, (s, \pi_s \circ \pi)) = m &
\end{aligned}
$$

It follows that σ^* is as required so that $\mathfrak{C}_1^*$ and $\mathfrak{C}_2^*$ are S-label variants. Since $\mathbf{R}$ satisfies Column Symmetry, it follows that for all $i \in N : \mathbf{R}(\mathfrak{C}_1^*, i) = \mathbf{R}(\mathfrak{C}_2^*, i)$. □

With Lemma 1 in place, proving the following result is now straightforward.

Theorem 1. *For all social choice rules* $\mathbf{S}$*, if* $\mathbf{S}$ *can be grounded in an original position argument, then* $\mathbf{S}$ *satisfies IISD.*

Proof. Suppose $\mathbf{S}$ is a social choice rule that can be grounded in an original position argument. Hence, there is some individual choice rule $\mathbf{R}$ such that for all scenarios $\mathfrak{C} = \langle N, A, S, d \rangle$ and all individuals $i \in N : \mathbf{R}(\mathfrak{C}^*, i) = \mathbf{S}(\mathfrak{C})$. Take two arbitrary Π-variants $\mathfrak{C}_1$ and $\mathfrak{C}_2$. Let $i \in N$ and $a \in A$ be arbitrary. We have

$a \in \mathbf{S}(\mathfrak{C}_1)$	
iff	Definition 14 (OP-argument)
$a \in \mathbf{R}(\mathfrak{C}_1^*, i)$	
iff	Lemma 1
$a \in \mathbf{R}(\mathfrak{C}_2^*, i)$	
iff	Definition 14 (OP-argument)
$a \in \mathbf{S}(\mathfrak{C}_2)$	

Hence, $\mathbf{S}$ satisfies IISD. □

Our next step is to show that if $\mathbf{S}$ can be grounded in an original position argument, it has to satisfy the IIPD axiom. We prove this in a way similar to before by showing that if two scenarios are Σ-variants, then if we apply any individual choice rule to their OP-transformations, we end up with the same set of admissible alternatives (Corollary 1). However, proving this claim requires a bit more preparatory work. Lemma 2 shows that the property of being Σ-variants is preserved under OP-transformations. Lemma 3 establishes that if two scenarios are Σ-variants, then the application of any individual choice rule on those scenarios themselves will yield the exact same recommendations. Corollary 1 follows immediately from these two properties.

Lemma 2. *For all choice scenarios* $\mathfrak{C}_1$ *and* $\mathfrak{C}_2$*: if* $\mathfrak{C}_1$ *and* $\mathfrak{C}_2$ *are* Σ*-variants then* $\mathfrak{C}_1^*$ *and* $\mathfrak{C}_2^*$ *are* Σ*-variants.*

Proof. Suppose that $\mathfrak{C}_1$ and $\mathfrak{C}_2$ Σ-variants. Hence, for all individuals $i \in N$ there is some bijection $\sigma_i : S_1 \to S_2$ such that for all alternatives $a \in A$: $d_1(i, a, s) = d_2(i, a, \sigma_i(s))$. For every $i \in N$ and state $(s, \pi) \in S^*$, let

$$\sigma_i^*(s, \pi) = (\sigma_{\pi(i)}(s), \pi)$$

Let $i \in N$, $a \in A$, $(s, \pi) \in S^*$, and $m \in \mathbb{R}$ be arbitrary. We have:

$$\begin{array}{rr} d_1^*(i, a, (s, \pi)) = m & \\ \text{iff} & \text{Definition 13 (OP-transformation)} \\ d_1(\pi(i), a, s) = m & \\ \text{iff} & \text{Definition 16 (Σ-variants)} \\ d_2(\pi(i), a, \sigma_{\pi(i)}(s)) = m & \\ \text{iff} & \text{Definition 13 (OP-transformation)} \\ d_2^*(i, a, (\sigma_{\pi(i)}(s), \pi) = m & \end{array}$$

Hence, for every $i \in N$, σ_i^* is as required so that $\mathfrak{C}_1^*$ and $\mathfrak{C}_2^*$ are Σ-variants. □

Lemma 3. *Let* $\mathbf{R}$ *be an individual choice rule. For all choice scenarios* $\mathfrak{C}_1 = \langle N, A, S, d_1 \rangle$ *and* $\mathfrak{C}_2 = \langle N, A, S, d_2 \rangle$*: if* $\mathfrak{C}_1$ *and* $\mathfrak{C}_2$ *are* Σ*-variants then for all* $i \in N : \mathbf{R}(\mathfrak{C}_1, i) = \mathbf{R}(\mathfrak{C}_2, i)$.

Proof. Let $\mathfrak{C}_1 = \langle N, A, S, d_1 \rangle$ and $\mathfrak{C}_2 = \langle N, A, S, d_2 \rangle$ be choice scenarios. Fix an arbitrary individual $i \in N$ and let $\mathbf{R}$ be an individual choice rule. Suppose that $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are Σ-variants. Hence, there is a bijection $\sigma_i : S \to S$ such that for all alternatives $a \in A : d_1(i, a, s) = d_2(i, a, \sigma_i(s))$. Let $\mathfrak{C}_i^\circ = \langle N, A, S, d_i^\circ \rangle$, where d_i° is defined as:

$$d_i^\circ(j, a, s) =_{df} \begin{cases} d_1(j, a, s) & \text{if j = i} \\ d_2(j, a, \sigma_i(s)) & \text{otherwise} \end{cases}$$

We show that $(\star)$ $\mathfrak{C}_i^\circ = \langle N, A, S, d_i^\circ \rangle$ and $\mathfrak{C}_2$ are S-label variants. Let $\sigma = \sigma_i$. Let $j \in N$, $a \in A$, $s \in S$, and $m \in \mathbb{R}$ be arbitrary. There are two cases. The case for $j \neq i$ follows immediately from the definition of d_i° and σ. If $j = i$, we have

$$\begin{array}{rr} d_i^\circ(i, a, s) = m & \\ \text{iff} & \text{Definition of } d_i^\circ \\ d_1(i, a, s) = m & \\ \text{iff} & \text{Definition of } \sigma \\ d_2(i, a, \sigma(s)) = m & \end{array}$$

Let $i \in N$, and $a \in A$ be arbitrary. We have:

$a \in \mathbf{R}(\mathfrak{C}_1, i)$

iff — Individualism and definition of d_i°

$a \in \mathbf{R}(\mathfrak{C}_i^\circ, i)$

iff — Column Symmetry and ($\star$)

$a \in \mathbf{R}(\mathfrak{C}_2, i)$

□

Corollary 1. *Let* $\mathbf{R}$ *be an individual choice rule. For all choice scenarios* $\mathfrak{C}_1 = \langle N, A, S, d_1\rangle$ *and* $\mathfrak{C}_2 = \langle N, A, S, d_2\rangle$*: if* $\mathfrak{C}_1$ *and* $\mathfrak{C}_2$ *are* Σ*-variants then for all* $i \in N : \mathbf{R}(\mathfrak{C}_1^*, i) = \mathbf{R}(\mathfrak{C}_2^*, i)$.

Theorem 2. *For all social choice rules* $\mathbf{S}$*, if* $\mathbf{S}$ *can be grounded in an original position argument, then* $\mathbf{S}$ *satisfies IIPD.*

Proof. The proof is analogous to the proof of Theorem 1 except that we rely on Corollary 1 instead of Lemma 1. □

4.2 Sufficient Conditions

In this section, we show that if a social choice rule $\mathbf{S}$ satisfies IISD and IIPD, then $\mathbf{S}$ can be grounded in an original position argument (Theorem 3). To prove this, we show that given any social choice rule $\mathbf{S}$ satisfying IISD and IIPD, we can define an individual choice rule $\mathbf{R}^{\mathbf{S}}_{=}$ such that $\mathbf{S}$ can be original position derived from $\mathbf{R}^{\mathbf{S}}_{=}$. In order to give a precise definition of $\mathbf{R}^{\mathbf{S}}_{=}$, we first introduce some additional notation.

Definition 17. *Let* $\mathfrak{C} = \langle N, A, S, d\rangle$ *be a choice scenario and let* $i \in N$*. The scenario* $\mathfrak{C}_{=i} = \langle N, A, S, d_{=i}\rangle$ *is such that for all* $j \in N$*,* $a \in A$*,* $s \in S$: $d_{=i}(j, a, s) = d(i, a, s)$.

In words, the scenario $\mathfrak{C}_{=i}$ is the scenario where at every outcome every individual receives the same payoff as individual i receives at that outcome in scenario $\mathfrak{C}$. To see how this works, an illustration is helpful (cf. Figure 7).

$\mathfrak{C}$	s_1	s_2
a	$(1,1)$	$(2,2)$
b	$(1,2)$	$(1,3)$
c	$(6,0)$	$(1,1)$

$\mathfrak{C}_{=1}$	s_1	s_2
a	$(1,1)$	$(2,2)$
b	$(1,1)$	$(1,1)$
c	$(6,6)$	$(1,1)$

Figure 7: A choice scenario ($\mathfrak{C}$) and the scenario ($\mathfrak{C}_{=1}$), where 1 is the "first" individual.

In the next step, we define the choice rule $\mathbf{R}^{\mathbf{S}}_{=}$ using the definition of $\mathfrak{C}_{=i}$ as follows.

Definition 18. *Let* $\mathbf{S}$ *be a social choice rule.* $\mathbf{R}^{\mathbf{S}}_{=}$ *is the choice rule such that for all choice scenarios* $\mathfrak{C} = \langle N, A, S, d\rangle$ *and all* $i \in N : \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}, i) = \mathbf{S}(\mathfrak{C}_{=i})$.

The first lemma that we establish shows that $\mathbf{R}^{\mathbf{S}}_{=}$ is an individual choice rule if $\mathbf{S}$ satisfies IIPD.

Lemma 4. *If* $\mathbf{S}$ *is a social choice rule satisfying IIPD, then* $\mathbf{R}^{\mathbf{S}}_{=}$ *satisfies:*

(i) Column Symmetry

(ii) Individualism

Proof. **Ad (i)**. Let $\mathfrak{C} = \langle N, A, S, d\rangle$ and $\mathfrak{C}' = \langle N, A, S', d'\rangle$ be S-label variants. Hence, there is a bijection $\sigma : S \to S'$ such that for all $i \in N$, $a \in A$: $d'(i, a, \sigma(s)) = d(i, a, s)$. Pick an arbitrary individual $i \in N$. That $\mathfrak{C}_{=i}$ and $\mathfrak{C}'_{=i}$ are S-variants is clear in view of the following:

1. For all $j \in N$: $d'_{=i}(j, a, \sigma(s)) = d'(i, a, \sigma(s))$ (Definition 17)
2. For all $j \in N$: $d_{=i}(j, a, s) = d(i, a, s)$ (Definition 17)
3. $d'(i, a, \sigma(s)) = d(i, a, s)$ (S-label variants)

Hence, for all $j \in N$ we have $d'_{=i}(j, a, \sigma(s)) = d_{=i}(j, a, s)$. Since $\mathbf{S}$ satisfies IIPD, it satisfies Column Symmetry (Fact 2), and so we have $\mathbf{S}(\mathfrak{C}_{=i}) = \mathbf{S}(\mathfrak{C}'_{=i})$. By Definition 18, $\mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}, i) = \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}', i)$.

Ad (ii). Let $\mathfrak{C} = \langle N, A, S, d\rangle$ and $\mathfrak{C}' = \langle N, A, S, d'\rangle$ be i-equivalent. Definition 17 implies that $\mathfrak{C}_{=i}$ and $\mathfrak{C}'_{=i}$ are identical scenarios. Since $\mathfrak{C}_{=i}$ and $\mathfrak{C}'_{=i}$ are identical, we have $\mathbf{S}(\mathfrak{C}_{=i}) = \mathbf{S}(\mathfrak{C}'_{=i})$. By Definition 18, $\mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}, i) = \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}', i)$. □

Before we come to the main result of this section, we prove two lemmas. First, if a social choice rule satisfies IISD, then the set of admissible alternatives is preserved under OP-transformations (Lemma 5). Second, for each individual i, the scenario $\mathfrak{C}^*_{=i}$ is a Σ-variant of $\mathfrak{C}^*$ (Lemma 6).

Lemma 5. *Let* $\mathbf{S}$ *be a social choice rule that satisfies IISD. For all choice scenarios* $\mathfrak{C}$*:* $\mathbf{S}(\mathfrak{C}) = \mathbf{S}(\mathfrak{C}^*)$.

Proof. Let $\mathfrak{C} = \langle N, A, S, d\rangle$ be a choice scenario and let $\mathbf{S}$ be a social choice rule that satisfies IISD. Where Π is the set of all bijections $\pi : N \to N$, $\mathfrak{C}^{\Pi} = \langle N, A, S^{\Pi}, d^{\Pi}\rangle$ is a State Multiplication variant of $\mathfrak{C}$ (Definition 11). Next, we show that $\mathfrak{C}^{\Pi}$ and $\mathfrak{C}^*$ are Π-variants. For every $i \in N$ and state $(s, \pi) \in S^{\Pi}$, let

$$\delta_{(s,\pi)}(i) = \pi^{-1}(i)$$

Let $i \in N$, $a \in A$, $(s, \pi) \in S^{\Pi}$, and $m \in \mathbb{R}$ be arbitrary. We have:

$$\begin{aligned}
d^{\Pi}(i, a, (s, \pi)) = m & \\
\text{iff} & \qquad \text{State Multiplication variants (Definition 11)} \\
d(i, a, s) = m & \\
\text{iff} & \qquad \text{OP-transformation (Definition 13)} \\
d^*(\pi^{-1}(i), a, (s, \pi)) = m & \\
\text{iff} & \qquad \delta_{(s,\pi)}(i) = \pi^{-1}(i) \\
d^*(\delta_{(s,\pi)}(i), a, (s, \pi)) = m &
\end{aligned}$$

Relying on our intermediate steps, we have the following, for every $a \in A$:

$$\begin{aligned}
a \in \mathbf{S}(\mathfrak{C}) & \\
\text{iff} & \qquad \text{State Multiplication Indifference} \\
a \in \mathbf{S}(\mathfrak{C}^{\Pi}) & \\
\text{iff} & \qquad \text{IISD} \\
a \in \mathbf{S}(\mathfrak{C}^*) &
\end{aligned}$$

□

Lemma 6. *For all choice scenarios $\mathfrak{C} = \langle N, A, S, d \rangle$ and $i \in N$: $\mathfrak{C}^*$ and $\mathfrak{C}^*_{=i}$ are Σ-variants.*

Proof. Fix a choice scenario $\mathfrak{C} = \langle N, A, S, d \rangle$ and individual $i \in N$. Because of the definition of Σ-variants (Definition 16) we must show that for all $j \in N$, there is a function $\sigma^j : S^* \to S^*$ such that for all $a \in A$ and $(s, \pi) \in S^*$:

$$d^*_{=i}(j, a, (s, \pi) = d^*(j, a, \sigma^j(s, \pi))$$

For every $j \in N$, we define σ^j as follows. Given some arbitrary state $(s, \pi) \in S^*$, let $\sigma^j(s, \pi) = (s, \pi')$, where π' is such that:

(1) $\pi(i) = \pi'(j)$

(2) $\pi(j) = \pi'(i)$

(3) $\pi(k) = \pi'(k)$ for all $k \in N \setminus \{1, j\}$

Note that σ^j is well-defined. First, there is always exactly one π' that satisfies conditions (1)-(3). Second, σ^j is a bijection; i.e if $\sigma^j(s, \pi) = (s, \pi')$, then $\sigma^j(s, \pi') = (s, \pi)$. As our next step, we show that we have the following property:

(†) for all $j \in N$, $a \in A$ and $(s, \pi) \in S^*$: $d^*(j, a, \sigma^j(s, \pi)) = d^*(i, a, (s, \pi))$

Let $j \in N$, $a \in A$, $(s, \pi) \in S^*$ be arbitrary and let $m \in \mathbb{R}$. Given (s, π), let π' be such that each of (1)-(3) is satisfied. We have:

$$\begin{array}{rr} d^*(j, a, \sigma^j(s, \pi)) = m & \\ \text{iff} & \text{by the definition of } \sigma^j \\ d^*(j, a, (s, \pi')) = m & \\ \text{iff} & \text{by condition (1)} \\ d^*(i, a, (s, \pi)) = m & \end{array}$$

By the definition of $d^*_{=i}$ (Definition 17): for all $j \in N$, $a \in A$ and $(s, \pi) \in S^*$: $d^*_{=i}(j, a, (s, \pi)) = d^*(i, a, (s, \pi))$. Hence, invoking our property (†), we conclude that for all $j \in N$, $a \in A$ and $(s, \pi) \in S^*$: $d^*_{=i}(j, a, (s, \pi)) = d^*(j, a, \sigma^j(s, \pi))$. □

Theorem 3. *Let* $\mathbf{S}$ *be a social choice rule. If* $\mathbf{S}$ *satisfies IISD and IIPD then* $\mathbf{S}$ *can be grounded in an original position argument.*

Proof. Let $\mathbf{S}$ be a social choice rule satisfying IISD and IIPD. We establish that for all choice scenarios $\mathfrak{C} = \langle N, A, S, d \rangle$ and all $i \in N$: $\mathbf{S}(\mathfrak{C}) = \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}^*, i)$, where $\mathbf{R}^{\mathbf{S}}_{=}$ is given by Definition 18. We can rely on Lemma 4 to establish that $\mathbf{R}^{\mathbf{S}}_{=}$ is an individual choice. Let $i \in N$ and $a \in A$ be arbitrary. We have

$$\begin{array}{rr} a \in \mathbf{S}(\mathfrak{C}) & \\ \text{iff} & \text{Lemma 5} \\ a \in \mathbf{S}(\mathfrak{C}^*) & \\ \text{iff} & \text{Lemma 6 and IIPD} \\ a \in \mathbf{S}(\mathfrak{C}^*_{=i}) & \\ \text{iff} & \text{Definition 18} \\ a \in \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}^*, i) & \end{array}$$

□

Corollary 2. *A social choice rule* $\mathbf{S}$ *can be grounded in an original position argument iff* $\mathbf{S}$ *satisfies IISD and IIPD.*

5 Concluding remarks

The main contribution of this paper is an axiomatization of the class of social choice rules that can be grounded in an original position argument within the context of choice under ignorance. In particular, we have shown that a social choice rule can be grounded in an original position argument if and only if the social choice rule satisfies the axioms of Indifference to Intra-State Distribution of Payoffs to Persons and Indifference to Intra-Person Distribution of Payoffs to States. Taken together, these axioms imply that it does not matter which individual gets what and under what circumstances.

While IISD and IIPD are rather strong, there are well-known social choice rules that satisfy them, the difference principle and the expected average utility rule being two examples. Notably, once we move to social choice rules that promote the welfare of those who are least well-off in some particular sense (e.g. maximin expected utility and difference expected utility), we find that these rules do not satisfy either IISD or IIPD and hence cannot be grounded in an original position argument. What conclusions can we draw from this?

If one takes original position arguments seriously, then our results can be used to criticize social choice rules on the grounds that they do not satisfy these two axioms. On that picture, the difference expected utility rule and the maximin expected utility rule are not fair after all.

We can also take the opposite perspective. If one finds the IISD and IIPD axioms unconvincing or if one insists that the difference expected utility rule or the maximin expected utility rule are plausible social choice rules, then original position arguments themselves must be considered unpersuasive.

Finally, one could take our results as a *reductio* against our formal explication of original position arguments and argue that there is something missing in our approach. We followed a standard approach to choice under ignorance. A consequence of this approach is that we treated ignorance in the original position in a one-dimensional way as ignorance about states *simpliciter*. One could also make a distinction between different types of ignorance in the original position: ignorance with respect to the identity of individuals and ignorance with respect to some independent state of nature. Hence, one option is to consider models that treat these two types of ignorance as separate from each other. In the absence of a unified theory of choice under multiple dimensions of ignorance, however, such a move could be *ad hoc*.

A Appendix

In this appendix, we state and prove two technical results that were referred to in Section 2.2.2. First, if we do not require Column Symmetry for individual choice, then original position arguments are trivial (Theorem 4). Second, if we require Column Symmetry but not Individualism, we have a characterization result with respect to IISD and Column Symmetry (Corollary 3).

Theorem 4. *Every social choice rule can be original position derived from a pointed choice rule that satisfies Individualism.*

Proof. Fix a social choice rule $\mathbf{S}$. Consider an arbitrary $\mathfrak{C} = \langle N, A, S, d\rangle$ and $i \in N$. We define $\mathbf{R}(\mathfrak{C}, i)$ by cases:

case 1: $\mathfrak{C}$ is such that $(\star)$ $S = S^\dagger \times \Pi$ and for all $\pi, \pi' \in \Pi$: if $\pi(i) = \pi'(i)$, then $d(i, a, (s, \pi)) = d(i, a, (s, \pi'))$. Let $\mathfrak{C}_i^\dagger = \langle N, A, S^\dagger, d_i^\dagger\rangle$ be such that for all $a \in A$, $j \in N$, and $s \in S^\dagger$: $d_i^\dagger(j, a, s) = d(i, a, (s, \pi))$ for some $\pi \in \Pi$ such that $\pi(i) = j$. Finally, let $\mathbf{R}(\mathfrak{C}, i) = \mathbf{S}(\mathfrak{C}_i^\dagger)$.

case 2: Condition ($\star$) does not apply. Then, let $\mathbf{R}(\mathfrak{C}, i) = A$.

Note first that $\mathbf{R}$ is well defined. To see why, it suffices to note that $\mathfrak{C}_i^\dagger$ is well defined whenever ($\star$) holds. Second, $\mathbf{R}$ satisfies Individualism: whether ($\star$) holds only depends on the payoffs of i, and if ($\star$) holds, then the definition of $\mathfrak{C}_i^\dagger$ only depends on the payoffs of i. Finally, we show that $\mathbf{S}(\mathfrak{C}) = \mathbf{R}(\mathfrak{C}^*, i)$ for every $i \in N$. To see why this holds, note that $(\mathfrak{C}^*)_i^\dagger = \mathfrak{C}$. So we have: $\mathbf{R}(\mathfrak{C}^*, i) = \mathbf{S}((\mathfrak{C}^*)_i^\dagger) = \mathbf{S}(\mathfrak{C})$. □

Theorem 5. *Let* $\mathbf{S}$ *be a social choice rule that satisfies IISD. There exists a pointed choice rule* $\mathbf{R}$ *such that (i)* $\mathbf{S}$ *can be original position derived from* $\mathbf{R}$ *and (ii)* $\mathbf{S}$ *satisfies CS iff* $\mathbf{R}$ *satisfies CS.*

Proof. Let $\mathbf{S}$ be a social choice rule that satisfies IISD. Let $\mathbf{R}$ be such that ($\star$) for all scenarios $\mathfrak{C} = \langle N, A, S, d\rangle$ and all $i \in N : \mathbf{R}(\mathfrak{C}, i) = \mathbf{S}(\mathfrak{C})$. Pick an arbitrary choice scenario $\mathfrak{C} = \langle N, A, S, d\rangle$. Since $\mathbf{S}$ satisfies SMI and IISD and by Lemma 5, $\mathbf{S}(\mathfrak{C}) = \mathbf{S}(\mathfrak{C}^*)$. Hence, given ($\star$) it follows immediately that for all $i \in N$: $\mathbf{S}(\mathfrak{C}) = \mathbf{R}(\mathfrak{C}^*, i)$, and that $\mathbf{R}$ satisfies CS if and only if $\mathbf{S}$ satisfies CS. □

Theorem 6. *Let* $\mathbf{S}$ *be a social choice rule. If* $\mathbf{S}$ *can be original position derived from a pointed choice rule* $\mathbf{R}$ *satisfying CS, then* $\mathbf{S}$ *satisfies CS and IISD.*

Proof. To show that $\mathbf{S}$ satisfies IISD, we can rely on Theorem 1 since its proof does not depend on Individualism. The only thing left to establish is that $\mathbf{S}$ satisfies CS. Suppose $\mathfrak{C}_1 = \langle N, A, S_1, d_1\rangle$ and $\mathfrak{C}_2 = \langle N, A, S_2, d_2\rangle$ are S-variants. Hence, there is some bijection $\sigma : S_1 \to S_2$ such that ($\star$) for all $i \in N$, $a \in A$, and $s \in S_1 : d_1(i, a, s) = d_2(i, a, \sigma(s))$. We show that if $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are S-variants, then $\mathfrak{C}_1^*$ and $\mathfrak{C}_2^*$ are S-variants. Let $\sigma^* : S_1^* \to S_2^*$ be such that for all $(s, \pi) \in S_1^* : \sigma^*(s, \pi) = (\sigma(s), \pi)$. First, σ^* is a bijection if σ is. Second, ($\star$) holds for σ^* as the required equalities are preserved by the OP-transformation. Hence, $\mathfrak{C}_1^*$ and $\mathfrak{C}_2^*$ are S-variants. Let $i \in N$, $a \in A$ be arbitrary. We have:

$a \in \mathbf{S}(\mathfrak{C}_1)$	
iff	Definition 14 (OP-argument)
$a \in \mathbf{R}(\mathfrak{C}_1^*, i)$	
iff	Column Symmetry for $\mathbf{R}$
$a \in \mathbf{R}(\mathfrak{C}_2^*, i)$	
iff	Definition 14 (OP-argument)
$a \in \mathbf{S}(\mathfrak{C}_2)$	

□

Corollary 3. *Let* **S** *be a social choice rule.* **S** *satisfies IISD and CS iff* **S** *can be original position derived from a pointed choice rule that satisfies CS.*

References

[1] Kenneth J Arrow and Leonid Hurwicz. An optimality criterion for decision-making under ignorance. In D F Carter and F Ford, editors, *Uncertainty and expectations in economics: Essays in honour of G. L. S. Shackle*. B. Blackwell, Oxford, 1972.

[2] Salvador Barbara and Matthew Jackson. Maximin, leximin, and the protective criterion: characterizations and comparisons. *Journal of economic theory*, 46(1):34–44, 1988.

[3] Ken Binmore. *Rational decisions*. Princeton University Press, Princeton NJ, 2008.

[4] Robert G Bone. Agreeing to fair process: The problem with contractarian theories of procedural fairness. *Boston University Law Review*, 83:485, 2003.

[5] Lara Buchak. Taking risks behind the veil of ignorance. *Ethics*, 127(3):610–644, 2017.

[6] Hun Chung. Rawls's self-defeat: A formal analysis. *Erkenntnis*, 85:1169–1197, 2020.

[7] Thijs De Coninck and Frederik Van De Putte. Original position arguments and social choice under ignorance. *Theory and decision*, pages 1–24, 2022.

[8] Gerald Gaus and John Thrasher. Rational choice and the original position: The (many) models of Rawls and Harsanyi. In Timothy Hinton, editor, *The original position*. Cambridge University Press, Cambridge, 2015.

[9] Johan E Gustafsson. Decisions under ignorance and the individuation of states of nature. Draft available at https://johanegustafsson.net/papers/decisions-under-ignorance-and-the-individuation-of-states-of-nature.pdf.

[10] Johan E Gustafsson. The difference principle would not be chosen behind the veil of ignorance. *The journal of philosophy*, 115(11):588–604, 2018.

[11] John C Harsanyi. Can the maximin principle serve as a basis for morality? A critique of John Rawls's theory. *American political science review*, 69(2):594–606, 1975.

[12] John C Harsanyi. *Rational behaviour and bargaining equilibrium in games and social situations*. Cambridge University Press, Cambridge, 1977.

[13] John Maynard Keynes. *A treatise on probability*. Macmillan, London, 1921.

[14] Will Kymlicka. *Contemporary political philosophy: An introduction*. Oxford: Oxford University Press, 2002.

[15] Eric Maskin. Decision-making under ignorance with implications for social choice. *Theory and decision*, pages 319–337, 1979.

[16] John Willard Milnor. Games against nature. In R M Thrall, C H Coombs, and R L Davis, editors, *Decision processes*. Wiley, New York, 1954.

[17] Michael Moehler. The Rawls–Harsanyi dispute: A moral point of view. *Pacific philosophical quarterly*, 99(1):82–99, 2018.

[18] Philippe Mongin and Marcus Pivato. Rawls's difference principle and maximin rule of allocation: A new analysis. *Economic theory*, 71(4):1499–1525, 2021.

[19] Juan D Moreno-Ternero and John E Roemer. The veil of ignorance violates priority. *Economics and philosophy*, 24(2):233, 2008.

[20] Derek Parfit. *Equality or priority. The Lindley Lecture.* University of Kansas, Department of Philosophy, 1991.

[21] Martin Peterson. *An introduction to decision theory.* Cambridge University Press, Cambridge, 2017.

[22] John Rawls. *A Theory of Justice.* Harvard University Press, Harvard, rev. ed. edition, 1999 (1971).

[23] Michael D Resnik. *Choices: An introduction to decision theory.* University of Minnesota Press, Minnesota, 1987.

[24] John E Roemer. Egalitarianism against the veil of ignorance. *The Journal of philosophy*, 99(4):167–184, 2002.

[25] Amartya Sen. *Collective choice and social welfare.* Holden Day, San Francisco, 1970.

[26] H Orri Stefánsson. Ambiguity aversion behind the veil of ignorance. *Synthese*, 198(7):6159–6182, 2021.

[27] Steven Strasnick. Social choice and the derivation of Rawls's difference principle. *The journal of philosophy*, 73(4):85–99, 1976.

[28] Philippe Van Parijs. Difference principles. In Samuel Freeman, editor, *The cambridge companion to John Rawls.* Cambridge University Press, Cambridge, 2001.