Economics & Philosophy (2025), **41**, 395–426
doi:10.1017/S0266267124000270

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Original position arguments: an axiomatic characterization

Thijs De Coninck[1,†] and Frederik Van De Putte[2,†]

[1]Independent Scholar, Belgium and [2]Erasmus Institute for Philosophy and Economics (EIPE), Erasmus School of Philosophy, Erasmus University Rotterdam, Rotterdam, the Netherlands
Corresponding author: Frederik Van De Putte; Email: vandeputte@esphil.eur.nl

(Received 18 October 2022; revised 18 June 2024; accepted 19 July 2024; first published online 02 December 2024)

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 supported by original position arguments. 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, against Rawlsian theories of procedural fairness, or in support of richer, multidimensional models of individual choice.

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' (Bone 2003: 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 (2003) 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

[†]Both authors contributed equally to this paper.
[1]While the terms 'rule' and 'procedure' have different meanings in more general contexts, we treat them as interchangeable in this paper.

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 this last approach is Rawls's original position argument in *A Theory of Justice* (Rawls 1999 [1971]).

The original position is a hypothetical situation where 'no one knows his place in society, his class position or social status' (Rawls 1999 [1971]: 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. (Rawls 1999 [1971]: 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 (Rawls 1999 [1971]: 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' (Rawls 1999 [1971]: 20).

The difference principle has been criticized by Sen (1970) 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 (1991) and Van Parijs (2001) 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 (1975, 1977) famously argued that when faced with complete ignorance, we should assign every outcome an equal probability and maximize expected utility.[4] 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 (Roemer 2002; Moreno-Ternero and

[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 De Coninck and Van De Putte (2023) for a critical survey of lexical variants of the difference principle from the viewpoint of original position arguments.

[4]In fact, already in a 1945 *Econometrica* paper, Vickrey proposed a similar approach to finding a fair welfare distribution, by 'choosing that distribution of income which such an individual would select were he asked which of various variants of the economy he would like to become a member of, assuming that once he selects a given economy with a given distribution of income he has an equal chance of landing in the shoes of each member of it' (Vickrey 1945: 329).

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

Roemer 2008; Gaus and Thrasher 2015; Buchak 2017; Moehler 2018; Gustafsson 2018; Chung 2020; Stefánsson 2021). Emerging from this is the view that, even if they are not able to single out a unique principle of justice, original position arguments do serve as a useful tool in sorting out intuitions regarding procedural fairness and its relation to social choice (Kymlicka 2002). What is lacking, however, is an exact and general characterization of which conceptions of justice or social choice can be supported by an original position argument. The present paper contributes to filling this gap.

1.1. 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 (Resnik 1987; Peterson 2017). 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 a class of social choice rules that can be supported by an original position argument.

We start by introducing our model of (social) choice under ignorance and define the general classes of individual and social choice rules that apply to this model (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 supported by 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 the normative implications of our results for original position arguments, arguing that they call for an enrichment of the uni-dimensional model of choice under ignorance that we assumed throughout (section 5).

1.2. Related work

Maskin (1979) 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 (1976) 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 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 De Coninck and Van De Putte (2023).

We use the same format, but make the underlying assumptions about individual and social choice rules fully precise in this paper. Moreover, whereas De Coninck and Van De Putte (2023) 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 supported by an original position argument.

2. Choice Scenarios and Choice Rules

We start by introducing a formal model of choice under ignorance (section 2.1). Once this is in place, we define and illustrate the notions of individual and social choice rules that take centre stage in this paper (section 2.2).

2.1. Choice scenarios and rules

The model we use is obtained by combining ingredients from the study of welfare distributions (Sen 1970) and decision-making under ignorance (Resnik 1987; Peterson 2017). A key ingredient is the class of choice scenarios, which serve as the input for individual and social choice.[5]

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.

A choice scenario is a compact representation of a situation where some agent – an individual member of society, a social planner or a group of persons – has to make a certain choice between several competing (mutually exclusive) alternatives. Here, N represents the set of individuals that make up society and A consists of the alternatives (options, choices, policies, courses of action) that may be chosen from. The set S represents the ignorance of the decision-maker, i.e. S is the set of states the decision-maker considers possible. Members of $A \times S$ are called the (possible) *outcomes* of the scenario. For each outcome (a, s), the distribution function d determines the welfare (level) $n \in \mathbb{R}$ of each individual $i \in N$ at s, given a. Note that the representation in terms of real numbers implies that welfare levels (of individuals, given states and alternatives) are interpersonally comparable and totally ordered. Depending on the application, one may also interpret them as *cardinal* welfare levels. In particular, some of the choice rules introduced below for illustrative purposes rely on notions of utility maximization, thus requiring such a stronger interpretation to make sense at all. However, none of our characterization results hinges on this. We work with finite models to keep our examples simple, but our characterization result also 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, so that individual 2 is

[5]Our notation and terminology follows De Coninck and Van De Putte (2023).

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

$\mathfrak{C}_1$	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.

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

Figure 2. Scenario $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are S-label variants, with $\sigma(s_1) = s'_2$ and $\sigma(s_2) = s'_1$.

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.

We consider two types of functions that are defined on the basis of choice scenarios: *social choice rules* and *individualistic choice rules*. Given any choice scenario $\mathfrak{C} = \langle N, A, S, d\rangle$, a social choice rule $\mathbf{S}$ determines a set of *socially admissible* alternatives $\mathbf{S}(\mathfrak{C}) \subseteq A$, whereas an individualistic choice rule $\mathbf{R}$ determines a set of *individually admissible* alternatives $\mathbf{R}(\mathfrak{C}, i) \subseteq A$ for each individual $i \in N$.[6] In addition, social and individualistic choice rules are supposed to satisfy some weak axioms that will be spelled out below. We will continue to use *choice rules* to refer to any rule that selects a set of admissible alternatives from the available ones, whether or not relative to some individual.

It should be noted that, while the specific individual and social choice rules that we will use to illustrate our results all maximize a certain social welfare ordering, our results themselves do not presuppose such an ordering at all. Put differently, the results concern *choice rules* in a very general sense, i.e. rules that select some (possibly empty) subset of the available options. All we require is that this selection is invariant under some minimal permutations of the models — as will become clear below.

The choice rules that we consider in this paper all satisfy Milnor's *Symmetry* condition (Milnor 1954), which means that they are invariant under any re-labelling of states or alternatives. In order to spell out this condition below, we define two equivalence relations on choice scenarios. First, two scenarios are S-label variants whenever they can be obtained from each other by a mere relabelling of the states. An example in case is given in Figure 2.

Definition 2 (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 $s \in S$: $d'(i, a, \sigma(s)) = d(i, a, s)$.*

[6]Note that we use the term 'admissible' here as a general denominator for those alternatives that are picked out by a given (individual or social) choice rule. This is not to be confused with the more narrow, game-theoretic reading of (weak) admissibility in terms of (weak) dominance.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

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

Figure 3. Scenario $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are A-label variants, with $\alpha(a) = c'$, $\alpha(b) = a'$, and $\alpha(c) = b'$.

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

$\mathfrak{C}_1^\Delta$	(s_1,δ_1)	(s_2,δ_1)	(s_1,δ_2)	(s_2,δ_2)
a	$(1,1)$	$(2,2)$	$(1,1)$	$(2,2)$
b	$(1,2)$	$(1,3)$	$(1,2)$	$(1,3)$
c	$(6,0)$	$(0,2)$	$(6,0)$	$(0,2)$

Figure 4. Scenario $\mathfrak{C}_1$ and $\mathfrak{C}_1^\Delta$, for $\Delta = \{\delta_1, \delta_2\}$.

Analogously, A-label variants are obtained by re-labelling the alternatives, holding all other parts of the scenario fixed. This is illustrated in Figure 3, where $\alpha(a) = c'$, $\alpha(b) = a'$, and $\alpha(c) = b'$. Formally:

Definition 3 (A-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* A-label variants *iff there exists a bijection $\alpha : A \to A'$ such that for all $i \in N$, $a \in A$ and $s \in S$: $d'(i, \alpha(a), s) = d(i, a, s)$.*

Finally, we will also presuppose that choice rules are invariant under the replication of states, as illustrated in Figure 4 and made precise by Definition 4.

Definition 4 (State Replication). *Let $\mathfrak{C} = \langle N, A, S, d\rangle$ be a choice scenario and let Δ be a non-empty finite set. Then $\mathfrak{C}^\Delta = \langle N, A, S^\Delta, d^\Delta\rangle$, where*

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

2.2. Choice rules

In what follows, we provide examples of individualistic (section 2.2.1) and social (section 2.2.3) choice rules and give exact definitions of the respective classes of such rules (sections 2.2.2 and 2.2.4). It should be emphasized that, while the concrete choice rules that we define below are well-known and have been defended in a range of contexts, we do not endorse or intend to defend them here; we merely use them for the didactic purpose of illustrating our formal framework and its implications.

2.2.1. Examples of individualistic 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.[7] More precisely, where $\mathsf{min}(X)$ denotes the $\leq$-minimal element of a set X of real numbers, we have:

Definition 5 (Maximin admissibility). *Where* $\mathfrak{C} = \langle N, A, S, d\rangle$ *is a choice scenario,* $i \in N$, *and* $a \in A$: $a \in \mathbf{R}^{\mathrm{m}}(\mathfrak{C}, i)$ *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.[8]

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, one may however rely on the *principle of insufficient reason* or *principle of indifference* (Keynes 1921), which states that in the absence of relevant evidence, individuals should assume that every state is equally likely.[9]

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 6 (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 $c \succ a \succ b$. 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.

[7]See Maskin (1979) for an axiomatic characterization of the maximin rule within the context of choice under ignorance.

[8]Maximin is often described as a conservative rule as it only takes into account the worst outcomes (Peterson 2017). To remedy this, a number of more sophisticated rules have been proposed, such as the leximin (Peterson 2017: §3.2) rule or lexical maximin rule (Resnik 1987: 27–28), and the optimism-pessimism rule (Resnik 1987: 32–35), also known as Hurwicz criterion after Hurwicz (1951) or alpha-index rule (Peterson 2017: §3.3). On the leximin rule, one first compares the worst outcomes of alternatives, but in case there is a tie, one looks at the second worst outcomes, and so on. On the optimism-pessimism rule, the value of an alternative a is given by a weighted average $\alpha \times \mathsf{max}(a) + (1 - \alpha) \times \mathsf{min}(a)$ of the best and the worst outcome for some *optimism index* $\alpha \in [0, 1]$. In this paper we stick to the simpler maximin rule, as we only use it for illustrative purposes.

[9]While it was endorsed by Vickrey and Harsanyi (cf. supra), the application of the principle of insufficient reason in the original position has been heavily criticized, not in the least by Rawls himself (Rawls 1999 [1971]). We remain neutral on this debate, and so we do not endorse said principle but merely use it to define various individual and social choice rules in what follows.

$\mathfrak{C}_1$	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 5. Scenario $\mathfrak{C}_1$ and $\mathfrak{C}_2$ are 1-equivalent but not 2-equivalent.

2.2.2. An exact characterization of individualistic choice rules

In general, we define individualistic choice rules as individual choice rules that satisfy certain axioms. Recall that by an individual choice rule we mean a choice rule that takes as input any choice scenario and individual, and gives as output a set of alternatives in the scenario that are admissible for that individual in that scenario. An individualistic choice rule is an individual choice rule that satisfies *Individualism*, *Symmetry* and *State Replication Indifference*. We spell out these three axioms in turn. 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. In order to define this axiom precisely, we introduce another equivalence relation on choice scenarios:

Definition 7 (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 $a \in A$ and all $s \in S$: $d'(i, a, s) = d(i, a, s)$.*

In Figure 5, 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.

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

As a second axiom, we presuppose the aforementioned *Symmetry* condition from Milnor (1954). This condition can be further subdivided into two parts: Column Symmetry requires that a choice rule is not sensitive to the way states are labelled; Row Symmetry requires that the label of alternatives is immaterial. Relying on the notions of invariance defined in section 2.1, we can state these conditions as follows:

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

Row Symmetry Let $\mathfrak{C} = \langle N, A, S, d\rangle$ and $\mathfrak{C}' = \langle N, A', S, d'\rangle$ be A-label variants, with $\alpha : A \to A'$ the underlying bijection. Then for all $i \in N$ and $a \in A$: $a \in \mathbf{R}(\mathfrak{C}, i)$ iff $\alpha(a) \in \mathbf{R}(\mathfrak{C}', i)$.

In what follows, Symmetry is simply the conjunction of both properties, Column Symmetry and Row Symmetry.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

Third and last, we require *State Replication Indifference*. This axiom is similar to Column Symmetry, but replacing S-label variants with state replication:

State Replication Indifference For all choice scenarios $\mathfrak{C} = \langle N, A, S, d \rangle$, all $i \in N$, and all non-empty and finite Δ: $\mathbf{R}(\mathfrak{C}, i) = \mathbf{R}(\mathfrak{C}^{\Delta}, i)$.

State Replication Indifference is weaker than the property known as Column Duplication (Milnor 1954; Binmore 2008).[10] The latter allows for adding single copies of just one state, whereas State Replication only allows one to copy all states a given number of times.

Let us discuss each of these three axioms in turn, in order to explain their motivation and intended interpretation. First, Individualism is integral to the concept of an original position argument. 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' (Rawls 1999 [1971]: 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.[11]

Second, Symmetry is both common and plausible for models that represent the ignorance of the decision-maker in unidimensional terms, i.e. as a single set of atomic states. Given such a representation, if the label of a state or alternative really matters for what counts as a *rational* choice, then it seems that something went wrong when stipulating the welfare levels in the first place. For instance, this may mean that depending on the labels, one should downgrade or upgrade the outcome (for certain individuals), associating a different payoff with the alternative at that state. Put differently, the whole point of representing a concrete situation by means of a decision scenario is that one must ensure that rational choice is really a function of the payoffs at each state, not of the labels or ordering of the states and alternatives themselves.[12] Moreover, as shown in the Appendix, without Column Symmetry, the notion of an original position argument as we characterize it is trivialized (Theorem 4).

Some readers may worry that our very starting point – i.e. that a decision-maker's ignorance can be represented by a unique set of atomic states (that cannot be further analysed) – is problematic in the context of original position arguments. After all, when placed in the original position, a decision-maker may well be ignorant about both the state of the world and their own identity. So even if in other contexts this is unproblematic, it would be a bad idea to collapse these two types of ignorance into one set of states. We have little to say in response to this critique, other than that it

[10]Column Duplication is also known as State-Individuation Invariance (Gustafsson 2022) or Independence of Duplicate States (Barbará and Jackson 1988).

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

[12]This is not to deny that as a matter of fact, people are often guided by the way alternatives and states are presented – a well-established fact in the literature on framing effects.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

requires spelling out in detail what would go wrong if we do collapse all ignorance into a single dimension or set of atomic states. Ultimately this is what our paper does: our results indicate that at least on some accounts of fair social choice rules, this model is overly simplistic. Our paper thus establishes a *conditional* claim: *if* one buys into the model defined here, *then* original position arguments can be characterized by the axioms that we discuss below. In section 5 we reconsider this point and discuss an alternative, richer account of ignorance.

Third and last, State Replication Indifference, though less well-known, seems just as natural as Symmetry in the context of choice under ignorance. Intuitively there should be no difference between scenarios such as $\mathfrak{C}_1$ and $\mathfrak{C}_1^{\Delta}$ in Figure 4. In particular, even if some bad, good or mediocre payoff will appear more often in the state replication variant of a given scenario than in the initial scenario, the ratio between such states and other states remains constant.[13]

In sum, we obtain the following:

Definition 8 (Individualistic choice rule). **R** *is an* individualistic choice rule *iff* **R** *is an individual choice rule that satisfies* Individualism, Symmetry *and* State Replication Indifference.

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 Symmetry and State Replication Indifference.[14] Thus, they are both individualistic choice rules in the above sense.

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' (Rawls 1999 [1971]: 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 about the state of the world, it is ambiguous what exactly 'least well-off' means, and hence one may specify this principle in conceptually distinct ways (see De Coninck and Van De Putte (2023) for an overview of these approaches).

Let us first focus on what is called the 'basic approach' in De Coninck and Van De Putte (2023). On the basic approach, we maximize welfare, ignoring the distinction between different states and different individuals.

[13]An analogous argument could be made for the uniform replication of alternatives. We will not go into the relevant axiom here, as it is irrelevant to our present results.

[14]In fact, Maximin satisfies the stronger Column Duplication property mentioned above. Expected Utility only satisfies State Replication Indifference.

Definition 9 (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* $b \in A : \mathsf{min}\{d(i, a, s) \,|\, i \in N \,\&\, s \in S\} \geq \mathsf{min}\{d(i, b, s) \,|\, i \in N \,\&\, s \in S\}$.

For example, in the scenario in Figure 1, we have $\mathsf{min}\{d(i, a, s) \,|\, i \in N \,\&\, s \in S\} = 1$, $\mathsf{min}\{d(i, b, s) \,|\, i \in N \,\&\, s \in S\} = 1$, and $\mathsf{min}\{d(i, c, s) \,|\, i \in N \,\&\, s \in S\} = 0$. 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 10 (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 $\mathrm{aeu}(a) = 1.5$, $\mathrm{aeu}(b) = 1.75$, and $\mathrm{aeu}(c) = 2$, 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 ($\mathrm{eu}_1(c) = 3$), whereas for individual 2 it is relatively low ($\mathrm{eu}_2(c) = 1$). 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.[15] 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 (2021).[16]

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.

[15]See however Chung (2023) for an argument against this criticism, in favour of utilitarianism.

[16]Mongin and Pivato (2021: 1504) 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.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

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

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

Definition 11 (Difference expected utility admissibility). *Where $\mathfrak{C} = \langle N, A, S, d\rangle$ is a choice scenario and $a \in A$: $a \in \mathbf{S}^{\mathrm{deu}}(\mathfrak{C})$ iff for all $b \in A$: $\mathsf{min}\{\mathsf{eu}_i(a) \mid i \in N\} \geq \mathsf{min}\{\mathsf{eu}_i(b) \mid i \in N\}$.*

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

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.*

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

Definition 12 (Maximin expected utility admissibility). *Where $\mathfrak{C} = \langle N, A, S, d\rangle$ is a choice scenario and $a \in A$, $a \in \mathbf{S}^{\mathrm{mau}}(\mathfrak{C})$ 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 $a \succ b \succ c$. An overview of the various rankings induced by the social choice rules introduced so far is given by Table 1.

2.2.4. An exact characterization of social choice rules

Recall that social choice rules are 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. As for individualistic choice rules, we require that social choice rules satisfy Symmetry and State Replication Indifference:[17]

Column Symmetry If $\mathfrak{C}$ and $\mathfrak{C}'$ are S-label variants then $\mathbf{S}(\mathfrak{C}) = \mathbf{S}(\mathfrak{C}')$.

[17]We use the same names as before, relying on the context to determine whether the individual or social version is meant.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

Row Symmetry Let $\mathfrak{C} = \langle N, A, S, d\rangle$ and $\mathfrak{C}' = \langle N, A', S, d'\rangle$ be A-label variants, with $\alpha : A \to A'$ the underlying bijection. Then for all $a \in A$: $a \in \mathbf{S}(\mathfrak{C})$ iff $\alpha(a) \in \mathbf{S}(\mathfrak{C}')$.

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

As before, we use *Symmetry* to refer to the conjunction of Column Symmetry and Row Symmetry. The interpretation of Symmetry and State Replication Indifference in the context of social choice rules is entirely analogous to that for individual choice rules (cf. section 2.2.2). Note that here, Symmetry does not say anything about the way distinct individuals are treated by a given social choice rule.

All the social choice rules introduced above satisfy these three axioms. For example, for the expected average utility rule, replicating 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 replication, the expected average utility also does not change. Likewise, none of the social choice rules we discussed refer to the labels of states or alternatives in determining which of those alternatives are admissible.

Recall that State Replication Indifference is weaker than Column Duplication (cf. section 2.2.2). While all four social choice rules introduced above satisfy State Replication Indifference, Column Duplication is only satisfied by the difference principle.[18]

Summing up the preceding, we obtain:

Definition 13 (Social choice rule). **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* **S** *satisfies Symmetry and State Replication Indifference.*

Our notion of social choice rules is very liberal: it assumes only very weak properties.[19] This is as intended. In what follows we ask, within this very broad class, which social choice rules can be supported by an original position argument.

3. Original Position Arguments

What does it mean that a given individualistic choice rule can be used to derive or support a social choice rule, using an original position argument? We start by giving

[18]To see why the difference principle satisfies Column Duplication, note that for this rule, whether a choice is admissible in some scenario only depends on the worst possible outcome, for one of the individuals in the scenario. Whether we duplicate some column will not make any difference to this. To see why average expected utility and both difference expected utility and maximin expected utility do not satisfy Column Duplication, a simple example suffices. Suppose that there are originally just two states s, s' and two options a, b. Let $d(i, a, s) = 3$, $d(i, b, s) = 2$, $d(i, a, s') = 1$, and $d(i, b, s') = 2$ for all $i \in N$. Note that since everyone's payoffs are identical, the three rules will give the same verdict. Then for each of the three rules, a and b will end up being equally good. In contrast, if we only duplicate state s', then b will end up being strictly better than a, again on each of the three rules.

[19]In particular, although this is the case for all the concrete examples that we use in this paper, we do not assume that social choice rules maximize some (social) ranking of alternatives.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

$\mathfrak{C}_1$	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 6. A choice scenario ($\mathfrak{C}_1$) and its OP-transformation ($\mathfrak{C}_1^*$).

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 supported by an original position argument (section 3.3).

3.1. A definition of original position arguments

We present the general format for evaluating original position arguments introduced in De Coninck and Van De Putte (2023). 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 initial choice scenario.

Definition 14 (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*

- $S^* = S \times \Pi$ and
- *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 6 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.

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

Definition 15 (Original position derivation). *Let* **S** *be a social choice rule, and let* **R** *be an individual choice rule.* **S** *can be original position derived from* **R** *iff for all scenarios* $\mathfrak{C} = \langle N, A, S, d\rangle$ *and all* $i \in N$: $\mathbf{S}(\mathfrak{C}) = \mathbf{R}(\mathfrak{C}^*, i)$.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

The social choice rule **S** *can be supported by an original position argument iff there is an individualistic choice rule* **R** *such that* **S** *can be original position derived from* **R**.

Note that, while social choice rules can be OP-derived from or supported by an (arbitrary) individual choice rules, in order to obtain an original position argument, we require the individual choice rule in question to satisfy the axioms of Symmetry, State Replication Indifference and Individualism. This allows us to consider what would happen to the notion of an original position argument if we lift some of these axioms (cf. our Appendix).

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 supported by 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 5] there is a $b \in A$ such that $\min\{d(i, b, s) \mid s \in S^*\} > \min\{d(i, a, s) \mid s \in S^*\}$ iff [By Fact 1] there is a $b \in A$ such that $\min\{d(j, b, s) \mid j \in N \,\&\, s \in S\} > \min\{d(j, a, s) \mid j \in N \,\&\, s \in S\}$ iff [By Definition 9] $a \notin \mathbf{S}^{\mathsf{d}}(\mathfrak{C})$. □

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 6] there is $b \in A$ such that $\mathsf{eu}_i(b) > \mathsf{eu}_i(a)$ iff [Notation 1 and simplifying the expression] $\sum[d(i, b, s) \,|\, s \in S^*] > \sum[d(i, a, s) \,|\, s \in S^*]$ iff [By Fact 1] there is $b \in A$ such that $\sum[d(j, b, s) \,|\, j \in N \,\&\, s \in S] > \sum[d(j, a, s) \,|\, j \in N \,\&\, s \in S]$ iff [Notation 2] there is $b \in A$ such that $\mathsf{aeu}(b) > \mathsf{aeu}(a)$ iff [By Definition 10] $a \notin \mathbf{S}^{\mathsf{aeu}}(\mathfrak{C})$. □

Proposition 3. *The difference expected utility principle cannot be supported by an original position argument.*

Proof. We show that there is no individualistic 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.[20]

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, $\mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_1) \neq \mathbf{S}(\mathfrak{C}_2)$.

The proof now proceeds by reductio. Suppose that there is some individualistic choice rule $\mathbf{R}$ such that the social choice rule $\mathbf{S}^{\mathsf{deu}}$ can be OP-derived from it. Hence, $x \in \mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_1)$ iff $x \in \mathbf{R}(\mathfrak{C}_1^*, i)$ for all $i \in N$. Similary for scenario $\mathfrak{C}_2$ we have $x \in \mathbf{S}^{\mathsf{deu}}(\mathfrak{C}_2)$ iff $x \in \mathbf{R}({\mathfrak{C}^*}_2, i)$ for all $i \in N$. By Column Symmetry and Individualism, for all $i \in N$, $x \in \mathbf{R}(\mathfrak{C}_1^*, i)$ iff $x \in \mathbf{R}(\mathfrak{C}_2^*, i)$. 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 supported by an original position argument (Propositions 1 and 2) as well as how one can argue that it is impossible to do so (Proposition 3).[21] 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 supported by an original position argument to the question of whether it satisfies these properties.

[20]Here, $\pi_=$ denotes the permutation that maps every individual onto itself, while $\pi_{\neq}$ maps the permutation that maps i to j and vice versa.

[21]One can similarly show that the maximin expected utility rule cannot be supported by an original position argument.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

$\mathfrak{C}_1$	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 7. Two Π variants.

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 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 16 (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 $\pi_s\colon N \to N$ such that for all $a \in A$, $i \in N$: $d'(\pi_s(i), a, s) = d(i, a, s)$.*

Figure 7 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 with 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: $\pi_s \neq \pi_{s'}$.

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 state 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 7, 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 $\mathfrak{C}_2$.

In light of the preceding, IISD is strictly stronger than *Anonymity*, i.e. the axiom that tells us that the labels of individuals do not matter (Sen 1970). Since we need not rely on *Anonymity* for our results to go through, we will omit a detailed discussion of this principle.

Our second axiom, *Indifference to Intra-Person Distribution of Payoffs at 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.

$\mathfrak{C}_1$	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 8. Two Σ-variants.

Definition 17 (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 8 depicts two Σ-variants. In scenario $\mathfrak{C}_2$, the payoffs of individual 1 are switched between s_1 and s_2 compared with 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 $\sigma_2(s_1) = s_1, \sigma_2(s_2) = s_2$.

Indifference to Intra-Person Distribution of Payoffs at 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 8. 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 for social choice rules. To see why, suppose $\mathfrak{C}$ and $\mathfrak{C}'$ are S-label variants (cf. Definition 2). 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 17). By IIPD, $\mathbf{S}(\mathfrak{C}) = \mathbf{S}(\mathfrak{C}')$. So we have:

Fact 2. *If* $\mathbf{S}$ *satisfies IIPD, then* $\mathbf{S}$ *satisfies Column Symmetry.*

Taken together, IISD and IIPD say that it does not matter which individual gets what and under what circumstances. More precisely, any permutation that swaps the payoffs for a given individual i at a state s with the payoffs of a (possibly different) individual j at a (possibly different) state s' – and that does so *uniformly* for each of the choices $a, b, \ldots$ – should leave the set of admissible choices unaltered. While these are rather strong conditions, both the basic difference rule and the

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

Table 2. Overview of the axioms governing social choice rules

	$\mathbf{s}^{d}$	$\mathbf{s}^{aeu}$	$\mathbf{s}^{deu}$	$\mathbf{s}^{mau}$
State Replication Indifference	+	+	+	+
Indifference to Intra-State Distribution of Payoffs to Persons	+	+	–	+
Indifference to Intra-Person Distribution of Payoffs at States	+	+	+	–

average expected utility rule satisfy them.[22] In contrast, as explained above, Difference Expected Utility does not satisfy IISD, while Maximin Expected Utility does not satisfy IIPD. Table 2 gives an overview of which social choice rules satisfy which of the axioms discussed so far.

4. Axiomatic Characterization

In this section, we prove that a social choice rule can be supported by an original position argument if and only if it satisfies Indifference to Intra-Person Distribution of Payoffs at 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. Left to right

In what follows we show that if a social choice rule can be supported by an original position argument, it satisfies IISD (Theorem 1) and IIPD (Theorem 2).

The following lemma establishes that if two scenarios are Π-variants, then for any individual, the same alternatives will be admissible in their OP-transformations. 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 individualistic 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 individualistic 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 16). 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:

$$d_1^*(i, a, (s, \pi)) = m$$

$$\text{iff} \qquad \text{Definition 14 (OP – transformation)}$$

[22]In fact, for these rules a yet stronger permutation invariance applies: for both rules, one may permute the payoffs for a given alternative (between individuals and states) *differently* from the payoffs for another alternative (cf. the role of Fact 1 in the proofs of Propositions 1 and 2).

$$
\begin{aligned}
& d_1(\pi(i), a, s) = m \\
& \qquad \text{iff} \qquad \text{Definition 16 } (\Pi - \text{variants}) \\
& d_2(\pi_s \circ \pi(i), a, s) = m \\
& \qquad \text{iff} \qquad \text{Definition 14 (OP} - \text{transformation)} \\
& d_2^*(i, a, (s, \pi_s \circ \pi)) = m.
\end{aligned}
$$

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 supported by an original position argument, then* $\mathbf{S}$ *satisfies IISD.*

Proof. Suppose $\mathbf{S}$ is a social choice rule that can be supported by an original position argument. Hence, there is some individualistic 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

$$
\begin{aligned}
& a \in \mathbf{S}(\mathfrak{C}_1) \\
& \qquad \text{iff} \qquad \text{Definition 15 } \big(\text{OP} - \text{argument}\big) \\
& a \in \mathbf{R}(\mathfrak{C}_1^*, i) \\
& \qquad \text{iff} \qquad \text{Lemma 1} \\
& a \in \mathbf{R}(\mathfrak{C}_2^*, i) \\
& \qquad \text{iff} \qquad \text{Definition 15 } \big(\text{OP} - \text{argument}\big) \\
& a \in \mathbf{S}(\mathfrak{C}_2).
\end{aligned}
$$

Hence, $\mathbf{S}$ satisfies IISD.

Our next step is to show that if $\mathbf{S}$ can be supported by an original position argument, it has to satisfy the IIPD axiom. We prove this in a way analogous to our proof of Theorem 1. More precisely, we show that if two scenarios are Σ-variants, then if we apply any individualistic 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 individualistic choice rule on those scenarios themselves will yield the exact same recommendations. Corollary 1 follows immediately from these two properties.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

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:

$$d_1^*(i, a, (s, \pi)) = m$$

iff Definition 14 (OP − transformation)

$$d_1(\pi(i), a, s) = m$$

iff Definition 17 (Σ − variants)

$$d_2\big(\pi(i), a, \sigma_{\pi(i)}(s)\big) = m$$

iff Definition 14 (OP − transformation)

$$d_2^*(i, a, \big(\sigma_{\pi(i)}(s), \pi\big) = m.$$

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* **R** *be an individualistic 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 **R** be an individualistic 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

$$d_i^\circ(i, a, s) = m$$

iff Definition of d_i°

$$d_1(i, a, s) = m$$

iff Definition of σ

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

$$d_2(i, a, \sigma(s)) = m.$$

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

$$\begin{aligned} & a \in \mathbf{R}(\mathfrak{C}_1, i) \\ & \qquad \text{iff} \qquad \text{Individualism and definition of } d_i^\circ \\ & a \in \mathbf{R}(\mathfrak{C}_i^\circ, i) \\ & \qquad \text{iff} \qquad \text{Column Symmetry and } (\star) \\ & a \in \mathbf{R}(\mathfrak{C}_2, i). \end{aligned}$$

Corollary 1. *Let* $\mathbf{R}$ *be an individualistic 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 supported by 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. Right to left

In this section, we show that if a social choice rule $\mathbf{S}$ satisfies IISD and IIPD, then $\mathbf{S}$ can be supported by 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 individualistic 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 18. *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 9).

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

Definition 19. *Let* $\mathbf{S}$ *be a social choice rule.* $\mathbf{R}^{\mathbf{S}}_{=}$ *is the individual 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 individualistic choice rule if $\mathbf{S}$ satisfies IIPD.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

$\mathfrak{C}$	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 9. A choice scenario ($\mathfrak{C}$) and the scenario ($\mathfrak{C}_{=1}$), where 1 is the 'first' individual.

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

(i) *Column Symmetry*
(ii) *Row Symmetry*
(iii) *State Replication Indifference*
(iv) *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 $i \in N$. Let $a \in A$ and $s \in S$. We have:

1. For all $j \in N$: $d_{=i}(j, a, s) = d(i, a, s)$ and $d'_{=i}(j, a, \sigma(s)) = d'(i, a, \sigma(s))$ (Definition 18)
2. $d'(i, a, \sigma(s)) = d(i, a, s)$ (S-label variants)

Hence, for all $j \in N$ we have $d'_{=i}(j, a, \sigma(s)) = d_{=i}(j, a, s)$. This implies that $\mathfrak{C}_{=i}$ and $\mathfrak{C}'_{=i}$ are S-label variants. Since $\mathbf{S}$ is a social choice rule, it satisfies Column Symmetry, and so we have $\mathbf{S}(\mathfrak{C}_{=i}) = \mathbf{S}(\mathfrak{C}'_{=i})$. By Definition 19, $\mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}, i) = \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}', i)$. Thus, $\mathbf{R}$ satisfies Column Symmetry.

Ad (ii). Let $\mathfrak{C} = \langle N, A, S, d\rangle$ and $\mathfrak{C}' = \langle N, A', S, d'\rangle$ be A-label variants. Hence, there is a bijection $\alpha : A \to A'$ such that for all $i \in N$, $s \in S$, and $a \in A$: $d'(i, \alpha(a), s) = d(i, a, s)$. Pick an arbitrary individual $i \in N$. Let $a \in A$ and $s \in S$. We have:

1. For all $j \in N$: $d_{=i}(j, a, s) = d(i, a, s)$ and $d'_{=i}(j, \alpha(a), s) = d'(i, \alpha(a), s)$ (Definition 18)
2. $d'(i, \alpha(a), s) = d(i, a, s)$ (A-label variants)

Hence, for all $j \in N$ we have $d'_{=i}(j, \alpha(a), s) = d_{=i}(j, a, s)$. This implies that $\mathfrak{C}_{=i}$ and $\mathfrak{C}'_{=i}$ are A-label variants. Since $\mathbf{S}$ is a social choice rule, it satisfies Row Symmetry, and so we have that for all $a \in A$, $a \in \mathbf{S}(\mathfrak{C}_{=i})$ iff $\alpha(a) \in \mathbf{S}(\mathfrak{C}'_{=i})$. By Definition 19, for every $i \in N$ and $a \in A$, $a \in \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}, i)$ iff $\alpha(a) \in \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}', i)$. Thus, $\mathbf{R}$ satisfies Row Symmetry.

Ad (iii). Analogous to the proof of item (i), substituting the notion of S-label variants with that of state replication variants.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

Ad (iv). Let $\mathfrak{C} = \langle N, A, S, d\rangle$ and $\mathfrak{C}' = \langle N, A, S, d'\rangle$ be i-equivalent. Definition 18 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 19, $\mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}, i) = \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}', i)$.

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 Replication variant of $\mathfrak{C}$ (Definition 4). We now 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:

$$\begin{aligned} d^\Pi(i, a, (s, \pi)) &= m \\ &\text{iff} && \text{State Replication Indifference} \\ d(i, a, s) &= m \\ &\text{iff} && \text{OP} - \text{transformation (Definition 14)} \\ d^*\left(\pi^{-1}(i), a, (s, \pi)\right) &= m \\ &\text{iff} && \delta_{(\mathrm{s},\pi)}(\mathrm{i}) = \pi^{-1}(\mathrm{i}) \\ d^*\left(\delta_{(s,\pi)}(i), a, (s, \pi)\right) &= m. \end{aligned}$$

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

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

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

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 17) 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 $(s, \pi) \in S^*$:

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

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) $\pi(j) = \pi'(i)$, and
(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 $\mathrm{j} \in \mathrm{N}, \mathrm{a} \in \mathrm{A}$ and $(\mathrm{s}, \pi) \in \mathrm{S}^*$: $d^*\big(j, a, \sigma^j(s, \pi)\big) = d^*(i, a, (s, \pi))$.

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:

$$d^*\big(j, a, \sigma^j(s, \pi)\big) = m$$

iff by the definition of σ^j

$$d^*\big(j, a, (s, \pi')\big) = m$$

iff by condition (1)

$$d^*(i, a, (s, \pi)) = m.$$

By the definition of $d^*_{=i}$ (Definition 18): 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* **S** *be a social choice rule. If* **S** *satisfies IISD and IIPD then* **S** *can be supported by an original position argument.*

Proof. Let **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 19. Note that by Lemma 4, $\mathbf{R}^{\mathbf{S}}_{=}$ is an individualistic choice rule. Let $i \in N$ and $a \in A$ be arbitrary. We have

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

$$
\begin{array}{rcc}
a \in \mathbf{S}(\mathfrak{C}) & & \\
 & \text{iff} & \text{Lemma 5} \\
a \in \mathbf{S}(\mathfrak{C}^*) & & \\
 & \text{iff} & \text{Lemma 6 and IIPD} \\
a \in \mathbf{S}(\mathfrak{C}^*_{=i}) & & \\
 & \text{iff} & \text{Definition 19} \\
a \in \mathbf{R}^{\mathbf{S}}_{=}(\mathfrak{C}^*, i). & &
\end{array}
$$

Corollary 2. *A social choice rule* **S** *can be supported by an original position argument iff* **S** *satisfies IISD and IIPD.*

5. Normative Implications of our Results

In the preceding, we saw that certain relatively simple social choice rules – i.e. difference expected utility and maximin expected utility – cannot be supported by an original position argument, at least not in the way we formalized the latter notion. Our axiomatization shows that this generalizes to any social choice rule that violates either IIPD or IISD.[23] This axiomatization result can in turn be used to choose between various prima facie defensible viewpoints on social choice rules and the normative relevance of original position arguments as we formalized them. In what follows, we discuss three such viewpoints in particular. The first one sticks to original position arguments (in the way we formalized them) as a crucial yardstick to determine the plausibility of social choice rules. If that is one's point of view, then one must reject rules such as difference expected utility and maximin expected utility, and more generally, any rule that violates IIPD or IISD. This still leaves one with a variety of possible social choice rules, depending on one's favoured individualistic choice rule (which may in turn depend on how one exactly conceives of the original position, cf. the Harsanyi vs. Rawls debate that we mentioned in the Introduction).

Second, one may deny the relevance of original position arguments for theories of (procedural) fairness. Our results can then be used to argue in favour of this viewpoint, insisting on the (putative) plausibility of rules like difference expected utility or maximin expected utility.

In fact, one need not even attach specific significance to any of those two rules to draw this conclusion. Consider the decision scenario $\mathfrak{C}_1$ depicted on the left-hand side of Figure 10.[24] In this scenario, we would expect that a is not socially admissible since it entails that the first individual is always worst off and the second always strictly better of. In addition, one may think that c is slightly better than b in that it always guarantees an equal payoff (*ex post*) for both individuals. However, for the

[23]Other examples of such rules can be found in De Coninck and Van De Putte (2023), based on a lexical interpretation of the difference principle.

[24]This decision scenario is partly based on Diamond (1967).

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

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

$\mathfrak{C}_3$	s_1	s_2
a	$(0,1)$	$(0,1)$
b	$(0,0)$	$(1,1)$
c	$(1,0)$	$(0,1)$

Figure 10. Three scenarios that are indistinguishable from the viewpoint of the original position.

(basic) Difference Principle and Average Expected Utility, all three options are equally good and hence all three of them are admissible.

This argument generalizes to any social choice rule that is supported by an original position argument. To see why, note that $\mathfrak{C}_1$ can be transformed into two other decision scenarios by switching the payoffs of both individuals at state s_1 (see scenario $\mathfrak{C}_2$ in the middle of Figure 10) or by switching the payoffs of the first individual across the two states (see scenario $\mathfrak{C}_3$ in the right-hand side of Figure 10). In other words, $\mathfrak{C}_2$ is a Π-variant of $\mathfrak{C}_1$, and $\mathfrak{C}_3$ is a Σ-variant of $\mathfrak{C}_1$. Relying on Row Symmetry, one can then easily observe that if any of the alternatives is admissible in any of these three scenarios, then all the alternatives must be admissible in all scenarios – at least according to any social choice rule that can be supported by an original position argument.

In sum, by requiring both IIPD and IISD of social choice rules, original position arguments obfuscate the distinction between individuals and states that is crucial in distinguishing the three alternatives in the above scenario. An obvious reaction to this would be to blame the axiom of Column Symmetry, since it implies that a social choice rule cannot look 'inside' the labels of states to notice e.g. that, given alternative a, individual i will have the worst payoff at any state, whereas given alternatives b and c, both individuals are worst off at some state.

So what if we simply drop Column Symmetry? It turns out that this trivializes our notion of original position arguments: *any* social choice rule can be OP-derived from an individual choice rule that merely satisfies Individualism. The reason is, roughly, that once the individual chooser is allowed to rely on the labels of states, it can simply 'notice' that the choice scenario it faces is the OP-transform of a different choice scenario $\mathfrak{C}$, and whenever that is the case, we may stipulate that the individual chooses whatever alternative would be chosen by the social choice rule in $\mathfrak{C}$ (see the Appendix to this paper for the full proof).

Regardless of said triviality, it seems strange to avoid the above problem simply by giving up Column Symmetry. As we argued above (cf. section 2.2.2), this axiom is quite fundamental for decision theory, as a theory of how rational individuals choose. The point seems not so much that labels of states should somehow become relevant (over and above the payoffs or utilities associated with states and alternatives), but rather that our ignorance should not be reduced to a single dimension or set of labels.

This then leaves us with the third defensible viewpoint, i.e. that our very notion of rational choice under ignorance needs to be enriched in order to account for the distinction between (ignorance about) individuals and states of the world. More precisely, if fairness requires that one makes those choices that an individually

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

rational agent would make in the original position, then that individual's choices must be represented in a more fine-grained way if we want it to be able to distinguish between a, b and c in the above example.

So what would the envisioned model look like? Although we must leave its full exploration for future work, let us briefly spell out the key idea behind it. In the most general terms, we can represent an individual's uncertainty not by a single set of states S, but by a set or tuple of primitive *dimensions of uncertainty*, $\mathbb{D} = \{D, D', \ldots\}$, treating the set of states $S = \times\mathbb{D}$ as a derived notion.[25] For instance, John's uncertainty about the weather may be distinct from his uncertainty about the actions of another rational agent such as Marie, and both may still be distinct from John's normative uncertainty – i.e. his uncertainty about how to best evaluate certain outcomes. When making rational decisions, we take into account the whole range of such uncertainties, but at the same time we (may) keep them apart in settling on some alternative.

One independent argument for representing such dimensions explicitly in a formal model is that different dimensions of uncertainty may come with different informational assumptions. For instance, John's uncertainty about the weather may be informed by some weather prediction app and thus be spelled out in terms of precise or imprecise probabilities; his normative uncertainty may derive from an ordinal ranking of normative theories; his uncertainty about what Marie will be doing could amount to plain ignorance.

While this is definitely a complication, such an enrichment allows us to both have our cake and eat it too. Although admissibility should be independent of the specific way one labels points on any dimension $D \in \mathbb{D}$, it need not be independent of the way welfare levels or payoffs are distributed across states more generally. For instance, John cannot readily swap the payoffs given

$$s = \langle \textsf{good weather}, \textsf{Marie goes to the cinema} \rangle$$

with those given

$$s' = \langle \textsf{bad weather}, \textsf{Marie does not go to the cinema} \rangle$$

even if he does so uniformly across possible alternatives. In contrast, if he uniformly swaps the labels $\textsf{bad weather}$ and $\textsf{good weather}$ throughout the scenario, this should not change his decision.

In the context of original position arguments, this enrichment means that the ignorance that was given in some initial decision scenario – say, one's uncertainty about the global economy – can be kept distinct from the ignorance about who one will end up being in society, in the OP-transformation of that scenario. This way one can support rules such as difference expected utility and maximin expected utility by an original position argument, and yet avoid that just any social choice rule can be so supported.

[25]Note again that we use 'uncertainty' here as an umbrella term for both ignorance and risk.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

6. Concluding Remarks

The main contribution of this paper is an axiomatization of the class of social choice rules that can be supported by an original position argument, within the context of choice under ignorance. In doing so, we followed a standard approach to individual choice under ignorance, modelling such ignorance in terms of a (single) set of possible states. Consequently, the ignorance in the original position transform of a decision scenario was represented by the Cartesian product of the ignorance in the initial scenario and the ignorance about one's identity in the society of individuals. Using this formal framework, we showed that a social choice rule can be supported by 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 at States. Taken together, these axioms imply that it does not matter which individual gets what and under what circumstances: as long as we do so uniformly across all alternatives, we can swap payoffs across individuals and across states.

As argued in section 5, this result implies that our model of original position arguments faces serious limitations. 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 supported by an original position argument. To solve this problem, a revision of the standard approach in terms of multiple dimensions of uncertainty is called for. Spelling out the formal details, potential applications and normative grounds for such a model is left for future work.

Acknowledgements. We wish to thank Marina Uzunova for helpful discussions that provided the basis of this paper and for comments throughout. Also thanks to Pawel Pawlowski and Joke Meheus for commenting on earlier drafts. Finally, we are indebted to three anonymous reviewers for incisive remarks on an earlier version. Thijs De Coninck's research was funded by the Research Foundation-Flanders (FWO Vlaanderen), grant no. 1167619N. Frederik Van De Putte's research was supported by the Research Foundation – Flanders (FWO-Vlaanderen), grant no. 12Q1918N, and by the project ENCODE: Explicating Norms of Collective Deliberation that is funded by the the Dutch Research Council (NWO), grant no. VI.Vidi.191.105.

References

Barbará S. and M. Jackson 1988. Maximin, leximin, and the protective criterion: characterizations and comparisons. *Journal of Economic Theory* **46**, 34–44.

Binmore K. 2008. *Rational Decisions*. Princeton: Princeton University Press.

Bone R.G. 2003. Agreeing to fair process: the problem with contractarian theories of procedural fairness. *Boston University Law Review* **83**, 485.

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

Chung H. 2020. Rawls's self-defeat: a formal analysis. *Erkenntnis* **85**, 1169–1197.

Chung H. 2023. When utilitarianism dominates justice as fairness: an economic defence of utilitarianism from the original position. *Economics & Philosophy* **39**, 308–333.

De Coninck T. and F. Van De Putte 2023. Original position arguments and social choice under ignorance. *Theory and Decision* **94**, 275–298.

Diamond P.A. 1967. Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment. *Journal of Political Economy* **75**, 765.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

Gaus G. and J. Thrasher 2015. Rational choice and the original position: the (many) models of Rawls and Harsanyi. In *The Original Position*, ed. T. Hinton. Cambridge: Cambridge University Press.

Gustafsson J.E. 2018. The difference principle would not be chosen behind the veil of ignorance. *Journal of Philosophy* **115**, 588–604.

Gustafsson J.E. 2022. Decisions under ignorance and the individuation of states of nature. *Thought: A Journal of Philosophy* **11**, 86–92.

Harsanyi J.C. 1975. Can the maximin principle serve as a basis for morality? A critique of John Rawls's theory. *American Political Science Review* **69**, 594–606.

Harsanyi J.C. 1977. *Rational Behaviour and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.

Hurwicz L. 1951. The Generalized Bayes Minimax Principle: A Criterion for Decision-Making Under Uncertainty. Cowless Commission Discussion Paper 355.

Keynes J.M. 1921. *A treatise on probability*. London: Macmillan.

Kymlicka W. 2002. *Contemporary Political Philosophy: An Introduction*. Oxford: Oxford University Press.

Maskin E. 1979. Decision-making under ignorance with implications for social choice. *Theory and Decision* **11**, 319–337.

Milnor J.W. 1954. Games against nature. In *Decision Processes*, ed. R.M. Thrall, C.H. Coombs and R.L. Davis. New York: Wiley.

Moehler M. 2018. The Rawls–Harsanyi dispute: a moral point of view. *Pacific Philosophical Quarterly* **99**, 82–99.

Mongin P. and M. Pivato 2021. Rawls's difference principle and maximin rule of allocation: a new analysis. *Economic Theory* **71**, 1499–1525.

Moreno-Ternero J.D. and J.E. Roemer 2008. The veil of ignorance violates priority. *Economics and Philosophy* **24**, 233–257.

Parfit D. 1991. *Equality or Priority. The Lindley Lecture*. Department of Philosophy, University of Kansas.

Peterson M. 2017. *An Introduction to Decision Theory*. Cambridge: Cambridge University Press.

Rawls J. 1999 [1971]. *A Theory of Justice. Revised edition*. Harvard: Harvard University Press,

Resnik M.D. 1987. *Choices: An Introduction to Decision Theory*. Minneapolis: University of Minnesota Press.

Roemer J.E. 2002. Egalitarianism against the veil of ignorance. *Journal of Philosophy* **99**, 167–184.

Sen A. 1970. *Collective Choice and Social Welfare*. San Francisco: Holden Day.

Stefánsson H.O. 2021. Ambiguity aversion behind the veil of ignorance. *Synthese* **198**, 6159–6182.

Strasnick S. 1976. Social choice and the derivation of Rawls's difference principle. *Journal of Philosophy* **73**, 85–99.

Van Parijs P. 2001. Difference principles. In *The Cambridge Companion to John Rawls*, ed. S. Freeman. Cambridge: Cambridge University Press.

Vickrey W. 1945. Measuring marginal utility by reactions to risk. *Econometrica* **13**, 319–333.

Appendix

In this appendix, we state and prove two technical results that were referred to in sections 2.2.2 and 5. 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 an individual 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 $\mathbf{R}(\mathfrak{C}, i) = \mathbf{S}(\mathfrak{C}_i^{\dagger})$.

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

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 an individual choice rule* $\mathbf{R}$ *such that (i)* $\mathbf{S}$ *can be original position derived from* $\mathbf{R}$ *and (ii)* $\mathbf{S}$ *satisfies Column Symmetry iff* $\mathbf{R}$ *satisfies Column Symmetry.*

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 Column Symmetry if and only if $\mathbf{S}$ satisfies Column Symmetry.

Theorem 6. *Let* $\mathbf{S}$ *be a social choice rule. If* $\mathbf{S}$ *can be original position derived from an individual choice rule* $\mathbf{R}$ *satisfying Column Symmetry, then* $\mathbf{S}$ *satisfies Column Symmetry 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 Column Symmetry. 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:

$$
\begin{array}{cc}
a \in \mathbf{S}(\mathfrak{C}_1) & \\
\text{iff} & \text{Definition 15 } \big(\text{OP} - \text{argument}\big) \\
a \in \mathbf{R}(\mathfrak{C}_1^*, i) & \\
\text{iff} & \text{Column Symmetry for } \mathbf{R} \\
a \in \mathbf{R}(\mathfrak{C}_2^*, i) & \\
\text{iff} & \text{Definition 15 } \big(\text{OP} - \text{argument}\big) \\
a \in \mathbf{S}(\mathfrak{C}_2). &
\end{array}
$$

Corollary 3. *Let* $\mathbf{S}$ *be a social choice rule.* $\mathbf{S}$ *satisfies IISD and Column Symmetry iff* $\mathbf{S}$ *can be original position derived from an individual choice rule that satisfies Column Symmetry.*

Thijs De Coninck was a PhD Student at the Centre for Logic and Philosophy of Science of Ghent University from 2018 till 2022 and predoctoral research fellow of the Flemish Research Foundation (FWO-Vlaanderen). He defended his PhD entitled *Lifting the veil: a formal analysis of original position arguments* in February 2023. Prior to defending his thesis, he had published on deontic logic (in the proceedings of the *DEON 2020/2021 conference*), on the philosophy of explanation (in *Synthese*), and on Rawlsian original position arguments (in *Theory and Decision*).

Frederik Van De Putte is Assistant Professor at the Erasmus Institute for Philosophy & Economics, Erasmus School of Philosophy, Erasmus University Rotterdam and guest professor at the Centre for Logic and Philosophy of Science of Ghent University. He obtained his PhD on the subject of prioritized adaptive

https://doi.org/10.1017/S0266267124000270 Published online by Cambridge University Press

logics (a specific type of non-monotonic logics) from Ghent University in May 2012 and has since published on a range of different subjects, including causal inference, the logic of (individual and collective) agency and obligation, modal logic more generally, and deontic logic. He is currently editor for *Erkenntnis* and is (co)editor of one edited book (on the contribution of Kit Fine to logic, edited with Springer) and two journal volumes. URL: https://frederikvandeputte.weebly.com/.

Cite this article: De Coninck T and Van De Putte F (2025). Original position arguments: an axiomatic characterization. *Economics and Philosophy* **41**, 395–426. https://doi.org/10.1017/S0266267124000270