



The limitations of the Arrovian consistency of domains with a fixed preference

James Nguyen^{1,2,3}

© The Author(s) 2019

Abstract

In this paper I investigate the properties of social welfare functions defined on domains where the preferences of one agent remain fixed. Such a domain is a degenerate case of those investigated, and proved Arrow consistent, by Sakai and Shimoji (Soc Choice Welf 26(3):435–445, 2006). Thus, they admit functions from them to a social preference that satisfy Arrow's conditions of Weak Pareto, Independence of Irrelevant Alternatives, and Non-dictatorship. However, I prove that according to any function that satisfies these conditions on such a domain, for any triple of alternatives, if the agent with the fixed preferences does not determine the social preference on any pair of them, then some other agent determines the social preference on the entire triple.

Keywords Social choice · Arrow · Fixed preferences · Restricted domains

1 Introduction and motivation

Social choice theory in the Arrovian style concerns functions from domains containing n -tuples of orders over some set of alternatives that map to an overall order on those alternatives. Arrow's conditions of Weak Pareto (WP), Independence of Irrelevant Alternatives (IIA), and Non-dictatorship (ND) put constraints on the behaviour these

I am particularly grateful to Nicholas Baigent and John Weymark for encouraging me to investigate the issues addressed in this paper and for reading it closely. Thanks also to Goretí Faria, Nicolas Wüthrich, and Alexandru Marcoci for comments on earlier drafts; Michael Morreau for stimulating discussions; and to audiences at the LSE Choice Group, the Bristol-LSE graduate workshop on formal epistemology, and the 2016 Central European Program in Economic Theory Annual Workshop in Udine for useful feedback.

✉ James Nguyen
james.nguyen@sas.ac.uk

¹ Department of Philosophy, University College London, London, UK

² Institute of Philosophy, School of Advanced Study, University of London, London, UK

³ Centre for Philosophy of Natural and Social Science, London School of Economics and Political Science, London, UK

functions, whilst the Universal Domain (UD) condition specifies their domain. Arrow (1963) proved that these conditions are not jointly satisfiable.

It is well known that UD is stronger than is needed to generate the impossibility result, but also well known that various restricted domains admit functions satisfying the remaining conditions (Gaertner 2001; Breton and Weymark 2011). In this paper, I investigate domains where the preferences of one individual are fixed, and the rest are allowed to vary. These domains are degenerate cases of domains containing what Sakai and Shimoji (2006) call a ‘dichotomous’ preference. An individual has a ‘dichotomous’ preference in their sense if the alternatives under consideration can be partitioned into two subsets, such that the individual in question prefers every element of one of the subsets over every element of the other, across the domain of the aggregation function. Clearly, if an individual has a fixed preference across the domain of the aggregation function, they have a dichotomous preference (unless of course their fixed preference is complete indifference). These domains do admit social welfare functions satisfying the remaining Arrovian conditions (Sakai and Shimoji 2006, Theorem 1), and thus, domains including a fixed preference are guaranteed to be Arrow consistent by the same result.

However, I prove here that according to any function on such a domain that satisfies WP, IIA, and ND, for any triple of alternatives, if the agent with the fixed preferences does not determine the social preference on any pair of them, then some other agent determines the social preference on the entire triple. Thus, the cost of accommodating varying social preferences over pairs of alternatives (varying in the sense of allowing disagreement with the agent with fixed preferences) is paid only by making some other agent determine the social preference over them across the domain. Equivalently, for any triple of alternatives, if the social ordering is to be sensitive to more than one agent’s preferences over them, then the fixed preference will always determine the social preference of at least one pair within the triple. This *prima facie* undesirable result suggests that the Arrovian consistency of such domains is guaranteed only at the cost of the preferences of the agent with the fixed preferences having significant influence on the social preference.

These domains are of interest for at least four reasons, each of them stemming from motivations for UD. When faced with a situation where an aggregation function is required I take it that, as a matter of fact, each individual has a specific, actual, preference ordering over the alternatives in question. Thus, there is only one actual preference profile to be aggregated.

However, Arrow (1963, Chapters III and VIII) motivated UD by noting that our *knowledge* about these preferences might be incomplete. If we cannot rule out any other profile this motivates UD.¹ In response to this we can consider contexts where the formal machinery of social choice theory is put to work in designing an actual aggregation procedure (such as in the literature on mechanism design). Presumably there exist scenarios where a social planner tasked with such a design problem may also be a member of the group of individuals whose preferences need to be aggregated. As long as she knows her own preferences in advance, then this undercuts the above

¹ In the first edition of *Social Choice and Individual Values* Arrow’s Condition 1 only required that there were at least three alternatives available such that every individual could order those three alternatives in anyway. The second edition’s Condition 1’ strengthened this and is now the standard UD assumption.

motivation for UD, because it requires that *every* individual may hold *any* possibly ordering on the alternatives. Domains where the social planner's preferences are held fixed, and the other individuals' preferences vary, correspond to the domains discussed in this paper.

An alternative motivation for UD that applies beyond mechanism design—including, for example, cases, where the machinery of social choice theory is used to investigate issues concerning justice—is offered by Kolm (1996) (whose views are usefully discussed by Weymark (2011)). Even if it is the case that in the actual world there is only once preference profile to be aggregated, whether or not the aggregation procedure is *just* depends on its counterfactual behaviour: what would have resulted had the individuals' preferences been otherwise. This corresponds to the behaviour of an aggregation function when applied to different preference profiles. If every individual involved *could* have had any preference over the alternatives in question, then this motivates UD. However, this might not always be the case. For some sets of alternatives, and for some individuals, their preferences may be such that they could not but hold the preferences that they do. The sorts of cases I have in mind include, for example, cases where the preferences of an individual are so central to their world view that they could not be otherwise without the individual failing to be that individual in question (this corresponds to the individual's preferences over the alternatives in question being one of what philosophers sometimes call their 'essential properties'). For example, for someone with strongly held religious beliefs, their preferences over some set of alternatives pertaining to those beliefs might be such that they simply could not hold different preferences without being a different person.

A third reason to think about these domains stems from a recent discussion in the philosophy of science literature. Okasha (2011) suggests that the question of how scientists should rationally choose between competing scientific theories (or models, or hypotheses) can be modelled in the social choice framework.² His argument is remarkably simple: take theoretical virtues—good-making features of scientific theories; like accuracy/fit-with-data, simplicity, and scope—to provide (ordinal) 'preference' rankings over the competing theories. The ranking by each virtue depends on how the competing theories compare with respect to that virtue. Then treat the question of theory choice as a question of coming up with a function from n -tuples of these orderings to an all-things-considered ranking.³ Crucially, Okasha suggests that the Arrovian conditions apply in such a context. If correct, Arrow's impossibility result rules out the possibility of rationally choosing between competing scientific theories in this manner.⁴

² I take it that his argument is supposed to be normative, not descriptive. Modelling the question of theory choice using the tools of social choice theory provides a way of asking what an ideally rational scientist should do when faced with competing scientific theories, rather than what an actual scientist does do.

³ As such the phrase 'theory choice' is something of a misnomer, since Okasha works in the framework of aggregating to an ordering rather than a choice function.

⁴ Okasha's paper has generated numerous responses and discussions (Morreau 2014, 2015; Rizza 2014; Okasha 2015; Stegenga 2015; Gaertner and Wüthrich 2016; Bradley 2016; Marcoci and Nguyen 2019). See also (Stegenga 2013; Cresto and Tajer 2019) for applications in the context of aggregating different sources of evidence.

Morreau (2015) argues against the applicability of UD in the theory choice context. If we again assume that there is a fact of the matter as to how each theoretical virtue ranks the competing theories, then again there is only one n -tuple to be aggregated. However, in analogy to Kolm's (1996) justification for UD in the social choice application, the *justification* for using a particular aggregation function depends on the behaviour of the function when augmented with different profiles, profiles of rankings *which might have been*. In addition, in the context of theory choice, at least some of the virtues, such as simplicity, might be unable to rank the competing alternatives in the way other than the one that they, in fact, do. As Okasha (2015, p. 288) puts it, simplicity ranks theories according to their 'essential features'. If, as a matter of fact, a given theory is simpler than a competing theory, then this is necessarily the case. Notice that this need not be the case for all theoretical virtues. Some, such as accuracy/fit-with-data, for example, could supply any ranking of the competing theories (where 'could' is read to indicate modal possibility). In his reply, Okasha admits that he is 'inclined to agree with Morreau's claim that some criteria of theory choice, such as simplicity, are rigid' (2015, p. 288).

If they are right about this, then the appropriate domain of a function from rankings of scientific theories by virtues is not UD, but one where at least one virtue—i.e., simplicity—supplies a fixed order on the alternatives, whilst the others—i.e., accuracy/fit-with-data—vary (I sharpen this in the next section). However, Okasha and Morreau disagree on the importance of this observation. Okasha claims that in 'the absence of a proof that [such a domain] is Arrow consistent, the right conclusion to draw is that we do not know whether an Arrovian impossibility result applies in this case or not. There is an unresolved mathematical question here' (2015, p. 290). As it happens, that such domains are Arrow consistent follows from the fact that an order fixed across such a domain is a degenerate case of a dichotomous order in the sense of (Sakai and Shimoji 2006), so the question has been resolved. However, the properties of the social welfare functions that satisfy the Arrovian conditions on such domains have not been investigated.

Finally, we can consider how to model non-welfarist, or more strictly speaking, mixed non-welfarist/welfarist, accounts of whether or not an aggregation function is just. Recall that on the standard social choice model, each element in an n -tuple of preferences to be aggregated is supposed to represent the preferences of an individual. Thus, whether or not an aggregation function is just depends solely on how it aggregates individuals' preferences involved (including what these might be in counterfactual situations). This way of addressing questions of social justice is strictly welfarist; the only relevant data concern the behaviour of an aggregation function when fed with individual preferences. However, there is nothing essential to the social choice framework that demands that the preferences within an n -tuple represent the preferences of an individual agent. We could instead take such an element to represent an objective comparison of the alternatives involved. If agreement with this objective comparison were to be the *sole* criterion on which to measure whether an aggregation function is just, as per a strictly non-welfarist account, then we would expect dictatorial (with respect to the objective comparison) functions to be just. However, mixed non-welfarist/welfarist accounts would allow considerations of the extent to which an aggregation function delivered the objective comparison of the alternatives involved

combined with the extent to which it matched the preferences of the individual agents involved. Depending on the nature of the alternatives in question, and the objective way of comparing them, it is plausible that the preference ranking representing the objective comparison of the alternatives would stay fixed irrespective of the subjective preferences of the individuals involved. Again, the domains of preferences considered in this paper correspond to such domains.

To aid readability I use the terminology associated with the standard application of the Arrovian framework throughout: I talk of ‘agents’, ‘individuals’, ‘preference orderings’, ‘indifference classes’, and so on. However, given the above considerations, it should be noted that the ‘preferences’ of an ‘agent’ or a ‘individual’ can represent the way a scientific virtue or a objective comparison orders the alternatives. I first recap Saki and Shimoji’s (2006) proof that the relevant domains admit functions satisfying IIA, ND, and WP. I then prove that provided there are enough alternatives, every agent can influence the value of such functions. I finally prove that according to any Arrow consistent function on such a domain, for any triple of alternatives, if the agent with the fixed preferences does not determine the social preference on any pair within the triple, then some other agent determines the social preference over the entire triple. Equivalently, for any triple of alternatives, if the social preference over them is not to be determined by a single agent across the domain, then the agent with the fixed preferences will determine the social preference over at least one pair from the triple across the domain. Thus, the cost of accommodating varying social preferences over pairs of alternatives is paid only by making agents with varying preferences determine the social preference. I return to how these formal results should be interpreted in each domain of application in the concluding section.

2 Formal Framework and Results

Let X denote a finite set of alternatives $X = \{a_1, a_2, \dots, a_m\}$, with variables x, y, z ranging over X . Let N denote a finite set of agents $N = \{1, 2, \dots, n\}$; \mathbf{P} be the set of strict orders over X ; and \mathbf{R} the set of weak orders over X . Each individual i is associated with orderings in \mathbf{R} . Let R_i and P_i denote the weak and strict parts of the preference order of an individual i , respectively. The domains of interest are $\{R_1\} \times \mathbf{R}^{n-1}$, $\{P_1\} \times \mathbf{R}^{n-1}$, and $\{P_1\} \times \mathbf{P}^{n-1}$, i.e., domains, where the preference order of agent 1 is fixed. Elements of these domains are profiles and are denoted by $\alpha, \beta, \gamma, \delta, \dots$. The restriction of a profile α to a pair of alternatives x, y is denoted by $\alpha|_{x,y}$. A social welfare function f is a function from a domain of profiles \mathfrak{D} to \mathbf{R} . When applied to a profile in \mathfrak{D} , the value of such a function is denoted by \mathcal{R} , whose strict part is denoted by \mathcal{P} . Where necessary, these are subscripted by Greek letters to indicate the particular profile they came from. The restriction of an ordering \mathcal{R} to a pair of alternatives x, y is denoted by $\mathcal{R}|_{x,y}$. Such a function $f : \mathfrak{D} \rightarrow \mathbf{R}$ satisfies the Arrovian conditions of Independence of Irrelevant Alternatives (IIA), Weak Pareto (WP), and Non-dictatorship (ND) if and only if:

- IIA** For all profiles $\alpha, \beta \in \mathfrak{D}$, and for all $x, y \in X$, if $\alpha|_{x,y} = \beta|_{x,y}$, then $(f(\alpha))|_{x,y} = (f(\beta))|_{x,y}$.

- WP** For all profiles in \mathcal{D} , and for all $x, y \in X$, if, for all $i \in N$, $x P_i y$, then $x \mathcal{P} y$.
- ND** There is no $i \in N$, such that for all profiles in \mathcal{D} , and all $x, y \in X$, if $x P_i y$, then $x \mathcal{P} y$.

For any domain \mathcal{D} , if there exists a $f : \mathcal{D} \rightarrow \mathbf{R}$, that satisfies IIA, WP, and ND, then \mathcal{D} is said to be *Arrow consistent*.

Theorem 1 Any domain of the form $\{R_1\} \times \mathbf{R}^{n-1}$ is Arrow consistent, for any $n \geq 2$, $|X| \geq 3$, and $R_1 \in \mathbf{R}$ with at least two indifference classes.

Proof This follows from Sakai and Shimoji (2006, Theorem 1) by noting that R_1 is a dichotomous preference. It is demonstrated by the below remark. □

Remark For any $n \geq 2$, $|X| \geq 3$, we can construct an $f : (\{R_1\} \times \mathbf{R}^{n-1}) \rightarrow \mathbf{R}$ which satisfies IIA, WP, and ND.⁵

Proof First notice that if R_1 is complete indifference over the alternatives in X , then individual 1 is a trivial dictator, violating ND. Therefore, we assume that R_1 contains at least two indifference classes. This means that we can find a partition of X with the following properties: $X_1 \cup X_2 = X$, $X_1 \cap X_2 = \emptyset$, for all $x, y \in X$, if $x \in X_1$, $y \in X_2$ then $x P_1 y$. Notice that since $|X| \geq 3$, then at least one of X_1, X_2 will contain at least 2 alternatives. Call that set X_i and the other X_j (no assumption is made about the cardinality of X_j other than the fact it is nonempty).

Then, we can define a function $f : (\{R_1\} \times \mathbf{R}^{n-1}) \rightarrow \mathbf{R}$ as follows:

- (i) For all $x, y \in X_i : x \mathcal{R} y \leftrightarrow x R_2 y$.
- (ii) For all $x, y \in X_j : x \mathcal{R} y \leftrightarrow x R_1 y$.
- (iii) For all $x \in X_i, y \in X_j : x \mathcal{R} y \leftrightarrow x R_1 y$.

f satisfies IIA, WP, and ND by construction. Moreover, it delivers a complete binary relation on X . It remains to show that it delivers a transitive relation. We prove this by reductio. Suppose there exists $x, y, z \in X$, such that $x \mathcal{R} y$, $y \mathcal{R} z$, and $z \mathcal{P} x$. The following situations are mutually exclusive and together exhaust the possibilities:

1. $x, y, z \in X_i$.
2. $x, y, z \in X_j$.
3. $x \in X_i, y, z \in X_j$.
4. $y \in X_i, x, z \in X_j$.
5. $z \in X_i, x, y \in X_j$.

⁵ This construction follows theirs, but was developed independently. It is a generalisation of a function suggested by Wulf Gaertner and John Weymark to Samir Okasha, cited in Okasha (2015, fn. 6). It builds on Fishburn’s (1976) discussion of ‘dictators on blocks’ of alternatives. Here, I restrict my focus to domains, where only one agent has a fixed preference and the rest of the agents’ preferences vary across all of \mathbf{R} . Sakai and Shimoji (2006) prove results with multiple fixed preferences. It bears noting here that their Theorem 2, which shows that in addition to satisfying IIA, WP, ND, there exist functions which satisfy ‘Anonymity’ (AN) (defined in the usual way by ensuring that social welfare functions act identically on permuted profiles) requires the existence of what they call an ‘isolated individual’, one whose preferences always disagree with any other individual’s (and thus, there does not exist a permutation on profiles, making AN vacuous), or that there exists a ‘dichotomous net’ on the domain, which requires for every triple of alternatives some agent with a fixed preference over those alternatives. These do not seem to match the motivations for considering a single fixed agent, as given in Sect. 1.

- 6. $x, y \in X_i, z \in X_j$.
- 7. $x, z \in X_i, y \in X_j$.
- 8. $y, z \in X_i, x \in X_j$.

By f 's conditions i and ii, and the fact that $X_i \cap X_j = \emptyset$, there will no intransitivity in cases 1 and 2. In cases 3 and 8 by the definitions of X_i and X_j , we know that either $x P_1 y, x P_1 z$ or $y P_1 x, z P_1 x$. By iii that means $x P z$ in the former case, and $y P x$ in the latter. Both contradict our reductio assumption. In cases 4 and 7 by the definitions of X_i and X_j , we know either $y P_1 x, y P_1 z$, or $x P_1 y, z P_1 y$. By iii that means $y P x$ in the former case, and $z P y$ in the latter. Both contradict our reductio assumption. In cases 5 and 6 by the definitions of X_i and X_j , we know either $z P_1 x, z P_1 y$, or $x P_1 z, y P_1 z$. By iii that means $z P y$ in the former case, and $x P z$ in the latter. Both contradict our reductio assumption. This exhausts all of the possible cases so f is guaranteed to deliver a transitive result. □

One might worry that no agent other than 1 or 2 has any effect on the value of the function. A way of capturing this worry is to place an additional constraint on such functions:

Liberalism (L) For every agent $i \in N$, there exists a pair $(x, y) \in X^2$, with $x \neq y$, such that if $x P_i y$, then $x P y$, and if $y P_i x$, then $y P x$ (Sen 1970, p. 153).

We can now prove that, providing that there are enough alternatives and that the agent, whose preferences are fixed strictly orders them, there exist functions that satisfy IIA, WP, and ND, which additionally satisfy L.

Theorem 2 For any $n \geq 2, |X| \geq 2n - 1$, and $P_1 \in \mathbf{P}$, there exists a function $f : (\{P_1\} \times \mathbf{R}^{n-1}) \rightarrow \mathbf{R}$ that satisfies IIA, WP, ND, and L.

Proof Assume there are n agents. Since $|X| \geq 2n - 1$ and the fixed agent provides a strict ordering⁶ we can partition X into n equivalence classes as follows: $\bigcup_{i=1}^n X_i = X$; for each $X_i, X_{j \neq i}, X_i \cap X_j = \emptyset; |X_1| = 1; |X_{i \neq 1}| = 2$; and for each X_i, X_j and $x, y \in X$, if $i > j$ and $x \in X_j$ and $y \in X_i$, then $x P_1 y$. This partition takes the top element of agents 1's ranking and places it in X_1 , the second and third elements are in X_2 , the fourth and fifth in X_3 , and so on, until reaching X_n , which will contain (at least 2 of) agent 1's least preferred alternatives. Now, define a function $f : (\{P_1\} \times \mathbf{R}^{n-1}) \rightarrow \mathbf{R}$ as follows:

- (i) For all X_i and $x, y \in X_i : x R y \leftrightarrow x R_i y$,
- (ii) For all $X_i, X_{j \neq i}$ and $x \in X_i$ and $y \in X_j : x R y \leftrightarrow x P_1 y$,

f satisfies IIA, WP, and ND by construction. Since the partition constructed contains n equivalence classes, i ensures that it satisfies L. As before, the value of f is complete and it remains to show that it is transitive. Suppose not. Then, there are some alternatives $x, y, z \in X$, such that $x R y, y R z$, and $z P x$. In this instance, it cannot be the case that all three are in the same equivalence class of the partition (since no

⁶ The below proof could be run using more alternatives and a fixed weak order, provided that it contained enough indifference classes with at least two elements.

$X_{i \neq n}$ contains more than two alternatives, and the social ordering over alternatives in X_n is guaranteed to be transitive by (i) and the fact that agent n supplies a transitive preference), so the following cases are mutually exclusive and jointly exhaustive:

1. There exists distinct X_i, X_j , such that $x \in X_i$ and $y, z \in X_j$.
2. There exists distinct X_i, X_j , such that $y \in X_i$ and $x, z \in X_j$.
3. There exists distinct X_i, X_j , such that $z \in X_i$ and $x, y \in X_j$.
4. There exists distinct X_i, X_j , and X_k , such that $x \in X_i, y \in X_j, z \in X_k$.

In case 4, by f 's condition ii, the social order of the alternatives will match agent 1's order over the alternatives, and since agent 1 provided a transitive preference order, there can be no intransitivity here. In case 1, suppose that $i > j$, then $y P_1 x$ by the construction of the partition and, therefore, $y \mathcal{P} x$ by f 's condition ii, contradicting the reduction assumption that $x \mathcal{R} y$. Alternatively suppose $j > i$, then $x P_1 z$ and $x \mathcal{P} z$, contradicting the reductio assumption that $z \mathcal{P} x$. In case 2, if $i > j$, then $z P_1 y$, so $z \mathcal{P} y$: contradiction. If $j > i$, then $y P_1 x$, so $y \mathcal{P} x$: contradiction. In case 3, if $i > j$, then $x P_1 z$, so $x \mathcal{P} z$ and if $j > i$, then $z P_1 y$, so $z \mathcal{P} y$ by the same reasoning. Both cases contradict the reductio assumption. Since this exhausts the cases, f cannot deliver an intransitive relation. \square

Theorems 1 and 2 show that by fixing the preference order by one agent both Arrow's impossibility result, and Sen's (1970) impossibility of the Paretian liberal result, can be avoided. However, they do so utilising aggregation functions that allow the fixed agent to determine the 'overall' structure of the aggregate order, and allowing other agents to 'fill in the gaps'. An interesting question is whether *any* Arrow consistent function on such a domain has to be this way. We can prove that something very like this has to be the case. To do so, it is useful to define the following notions: an agent i is *weakly decisive* on an ordered pair $(x, y) \in X^2$ if and only if in all profiles, where $x P_i y, x \mathcal{P} y$. An agent who is weakly decisive on (x, y) and (y, x) is said to be *decisive* on $\{x, y\}$. Finally, an agent i is *pivotal* at a profile $\alpha = (R_1, \dots, R_m)$ if and only if there exists a profile $\beta = (R'_1, \dots, R'_m)$, such that $R_j = R'_j$ for all $j \neq i, R_i \neq R'_i$, and $f(\alpha) \neq f(\beta)$ (Barberá 1980).

Theorem 3 For any $n \geq 2, |X| \geq 3$, and $P_1 \in \mathbf{P}$, and function $f : (\{P_1\} \times \mathbf{P}^{n-1}) \rightarrow \mathbf{R}$ that satisfies IIA and WP: for any triple of distinct alternatives $x, y, z \in X$, if agent 1 is not decisive on any pair from those alternatives, then there exists an agent who is decisive on all of them.

Proof We prove the theorem by assuming that f satisfies IIA and WP. Without loss of generality choose a triple of distinct alternatives $x, y, z \in X$ such that $x P_1 y P_1 z$. We assume that agent 1 is not decisive on any pair of those alternatives to show that some agent i is decisive on all of them. The proof proceeds in six steps.⁷ Step 1 shows that there exists a pivotal agent i at some profile β . Step 2 shows that i is weakly decisive on the pair (x, y) . Step 3 shows that i is weakly decisive on (y, x) , thereby establishing that i is decisive on $\{x, y\}$. Step 4 shows that i is also decisive on $\{x, z\}$. Step 5 shows that the same process delivers an agent j who is decisive on $\{y, z\}$. Step 6 shows that $i = j$.

⁷ The strategy in steps 1–5 is similar to Yu's (2012) proof of Arrow's Theorem, which is in turn based upon the third proof from Geanakoplos (2005).

Step 1. As agent 1 is not decisive on any pair from $\{x, y, z\}$, and 1 is trivially weakly decisive on the pair (z, y) , we know that there exists a profile α , such that $z\mathcal{R}y$. Order the agents, as displayed below (Fig. 1). Now, one by one from the right of the line onwards, perform a pairwise reversal on the preferences over y and z for each agent who prefers z to y . By WP, at some point in this process, we will arrive at a profile, where $y\mathcal{P}z$, and where no further pairwise reversals of the preferences of agents later in the sequence will change this. Call the agent who induced the (final) switch i . This provides the profiles β and γ (Figs. 2, 3). This suffices to show that i is pivotal at β (and γ).

Step 2. We now show that i is weakly decisive on (x, y) . Insert x directly above y in γ for each agent $\neq 1$, and for 1 we know that xP_1yP_1z . This delivers a profile δ (Fig. 4). By WP, $x\mathcal{P}_\delta y$. By IIA and the fact that $\delta|_{y,z} = \gamma|_{y,z}$, $y\mathcal{P}_\delta z$. So by transitivity we

Fig. 1 $\alpha : z\mathcal{R}y$

1
...
y	...	y	z ... z
...
z	...	z	y ... y
...

Fig. 2 $\beta : z\mathcal{R}y$

1	...	i	...
...
y	...	y	z z ... z
...
z	...	z	y y ... y
...

Fig. 3 $\gamma : y\mathcal{P}z$

1	...	i	...
...
y	...	y	y z ... z
...
z	...	z	z y ... y
...

Fig. 4 $\delta : x\mathcal{P}y\mathcal{P}z$

1	...	i	...
...
...
x
...	...	x	x ...
y	...	y	y z ... z
...
...	x ... x
z	...	z	z y ... y
...
...

have $x \mathcal{P}_\delta z$. Now, rearrange the preferences over x, y of every agent $\neq i, 1$ as you like, keeping their relative position with respect to z fixed. Raise z in i 's preference ranking, such that: $x P_i z P_i y$. This delivers the profile ϵ where each ? denotes a possible location for x (Fig. 5). By IIA and $\epsilon|_{y,z} = \beta|_{y,z}, z \mathcal{R}_\epsilon y$. By IIA and $\epsilon|_{x,z} = \delta|_{x,z}, x \mathcal{P}_\epsilon z$. By transitivity, $x \mathcal{P} y$. This holds for every way that x can be distributed across the question marks, which exhausts all relative orders of x and y (where $x P_i y$). By IIA, this makes i weakly decisive on (x, y) .

Step 3. We now show that i is weakly decisive on (y, x) and, therefore, decisive on $\{x, y\}$. Since agent 1 is assumed not to be decisive on $\{x, y\}$ and is trivially weakly decisive on (y, x) , there must exist a profile ζ , such that $y \mathcal{R}_\zeta x$. Distribute x around (i.e., directly above or below) y in β for all agents $\neq 1$, so the relative ordering of x and y in the result matches ζ . This gives the profile η , where the † indicates a possible location for x (Fig. 6). By IIA and $\zeta|_{y,x} = \eta|_{y,x}, y \mathcal{R}_\eta x$. By IIA $\beta|_{y,z} = \eta|_{y,z}, z \mathcal{R}_\eta y$. By transitivity $z \mathcal{R}_\eta x$.

Now, raise y above z in i 's preference ranking and rearrange the relative ranking of x and y as you like for all agents other than 1 and i . This delivers the profile θ (Fig. 7). By IIA and $\theta|_{y,z} = \gamma|_{y,z}, y \mathcal{P}_\theta z$. By IIA and $\theta|_{x,z} = \eta|_{x,z}, z \mathcal{P}_\theta x$. By transitivity $y \mathcal{P}_\theta x$. This holds for every way that x can be distributed across the question marks, which exhausts all relative orders of x and y (where $y P_i x$). By IIA, this makes i weakly decisive on (y, x) . Which combined with the fact that i is weakly decisive on (x, y) makes i decisive on $\{y, x\}$.

Fig. 5 $\epsilon : x \mathcal{P} z \mathcal{P} y$

1	...		i
...
x
...	...	?	x
y	...	y	z	z	...	z
...	...	?	y
...
...	?	...	?
z	...	z	...	y	...	y
...	?	...	?
...

Fig. 6 $\eta : z \mathcal{R} y \mathcal{R} x$

1	...		i
...
x
...	...	†
...	...	y	z	z	...	z
y	...	†
...
...	†	†	...	†
z	...	z	y	y	...	y
...	†	†	...	†
...

Fig. 7 $\theta : y\mathcal{P}_z\mathcal{R}_x$

1	...		i	...		
...
x
...	...	?	y
...	...	y	z	z	...	z
y	...	?
...
...	?	?	...	?
z	...	z	?	y	...	y
...	?	...	?
...

Fig. 8 $\iota : x\mathcal{P}_z\mathcal{R}_y$

1	...		i	...		
...
...	x	x	...	x
x	...	y	z	z	...	z
...
...
y
...
...	...	x
z	...	z	y	y	...	y
...

Fig. 9 $\kappa : x\mathcal{P}_y\mathcal{P}_z$

1	...		i	...		
...
...	x	?	...	?
x	...	y	y	z	...	z
...	z	?	...	?
...
y
...
...	...	?
z	...	z	...	y	...	y
...	...	?

Step 4. We now show that i is decisive on $\{x, z\}$ using the same technique as above. For all agents $\neq 1$ insert x directly above z in β to deliver the profile ι (Fig. 8). Since $\iota|_{z,y} = \beta|_{z,y}, z\mathcal{R}_iy$. By WP, $x\mathcal{P}_iz$. By transitivity $x\mathcal{P}_iy$. Now, raise y above z in i 's preference ranking and rearrange the relative ranking of x and z as you like for all other than 1 and i . This delivers the profile κ (Fig. 9). By IIA and $\kappa|_{x,y} = \iota|_{x,y}, x\mathcal{P}_\kappa y$. By IIA and $\kappa|_{y,z} = \gamma|_{y,z}, y\mathcal{P}_\kappa z$. By transitivity $x\mathcal{P}_\kappa z$. This holds for each way that x can be distributed across the question marks, which exhausts all relative orders of x and z (where $x\mathcal{P}_iz$). By IIA this makes i weakly decisive on (x, z) .

Fig. 10 $\mu : y\mathcal{P}_z\mathcal{R}_\lambda x$

	1	...		i
...
...	†	...	†
x	...	y	y	z	...	z	
...	†	...	†
...
y	...	†	†	y	...	y	
...	...	z	z	
...	...	†	†	
z	
...	

Fig. 11 $v : z\mathcal{R}_\lambda y\mathcal{P}_\mu z$

	1	...		i
...
...	z	?	...	?	
x	...	y	y	z	...	z	
...	?	...	?	
...	
y	...	?	...	y	...	y	
...	...	z	?	
...	...	?	?	
z	
...	

We now show that i is weakly decisive on (z, x) . Since we are assuming that agent 1 is not decisive on any pair from $\{x, y, z\}$, and 1 is trivially weakly decisive on (z, x) , to avoid 1 being decisive on $\{x, z\}$, there must exist a profile λ , where $z\mathcal{R}_\lambda x$. Distribute x directly above or below z in γ to match the relative order of x and y in λ which delivers a profile μ , where the † denotes a possible location for x (Fig. 10). By IIA and $\lambda|_{z,x} = \mu|_{z,x}$, $z\mathcal{R}_\mu x$. By IIA and $\gamma|_{z,y} = \mu|_{z,y}$, $y\mathcal{P}_\mu z$. By transitivity $y\mathcal{P}_\mu x$. Now, raise z above y in i 's preference ranking and rearrange the relative ranking of x and z as you like for all other that 1 and i . This delivers the profile v (Fig. 11). By IIA and $v|_{x,y} = \mu|_{x,y}$, $y\mathcal{P}_v x$. By IIA and $v|_{y,z} = \beta|_{y,z}$, $z\mathcal{R}_v y$. By transitivity $z\mathcal{P}_v x$. This holds for each way that x can be distributed across the question marks, which exhausts all relative orders of x and z (where $zP_i x$). By IIA this makes i weakly decisive on (z, x) , which combined with the previous result means that i is decisive on $\{z, x\}$.

Step 5. We now show that there is an agent j who is decisive on $\{y, z\}$. We proceed in similar manner to the above. First, to avoid agent 1 being decisive on $\{y, x\}$, there must exist a profile, where $y\mathcal{R}_\lambda x$. Order the agents, so that those to the left prefer x to y , and those to the right prefer y to x . Perform the same pairwise switch process as occurred in step 1. By WP, we know that at some point in this process, we reach $x\mathcal{P}_\lambda y$, where no further pairwise reversals later in the sequence will change this. This provides the profiles ξ and o (Figs. 12, 13). Insert z directly below y for all $\neq 1$ in o , and we know that $yP_1 z$. This delivers the profile π (Fig. 14), such that $\pi|_{x,y} = o|_{x,y}$, so by IIA $x\mathcal{P}_\pi y$, and $y\mathcal{P}_\pi z$ by WP, so $x\mathcal{P}_\pi z$ by transitivity. Now, keeping their relative

Fig. 12 $\xi : y\mathcal{R}x$

1	...	j	...
...
x	...	x	y
...
y	...	y	x
...

Fig. 13 $o : x\mathcal{P}y$

1	...	j	...
...
x	...	x	y
...
y	...	y	x
...

Fig. 14 $\pi : x\mathcal{P}y\mathcal{P}z$

1	...	j	...
...
x	...	x	y
...	z
y	...	y	x
...	...	z	z
...
z
...

Fig. 15 $\mu : y\mathcal{R}x\mathcal{P}z$

1	...	j	...
...
x	...	x	y
...	x
y	...	y	z
...	...	z	x
...
z
...

ranks with respect to x fixed, arbitrarily rearrange the ranking of y, z for every agent (other than j and 1), and raise y over x in j 's ranking. This delivers the profile μ again the ?s denote possible locations of z (Fig. 15). By IIA and $\mu|_{x,y} = \xi|_{x,y}, y\mathcal{R}_\mu x$. By IIA and $\mu|_{x,z} = \pi|_{x,z}, x\mathcal{P}_\mu z$. So by transitivity, $y\mathcal{P}_\mu z$. This holds for every y that z can be distributed across the question marks, which exhausts all relative orders of z and y (where $y\mathcal{P}z$). By IIA, this makes j weakly decisive on (y, z) .

We now show that there exists a agent $k \neq 1$ who is weakly decisive on (z, y) . Since agent 1 is trivially weakly decisive on (z, x) , to avoid agent 1 being decisive on $\{z, x\}$, there must exist a profile σ , where $z\mathcal{R}_\sigma x$. Perform the same ordering procedure and pairwise switching of x and z , and by WP, we know that the two profiles τ and ν exist (Figs. 16, 17). Since agent 1 is trivially weakly decisive on (z, y) to avoid agent 1

Fig. 16 $\tau : z\mathcal{R}x$

	1	...	k	...
...
x	...	x	z	z ... z
...
z	...	z	x	x ... x
...

Fig. 17 $v : x\mathcal{P}z$

	1	...	k	...
...
x	...	x	x	z ... z
...
z	...	z	z	x ... x
...

Fig. 18 $\chi : x\mathcal{P}z\mathcal{R}y$

	1	...	k	...
...
x	...	x	x	\dagger ... \dagger
...	z ... z
...	\dagger ... \dagger
y
...	...	\dagger	\dagger	x ... x
...	...	z	z	...
z	...	\dagger	\dagger	...
...

Fig. 19 $\psi : z\mathcal{R}x\mathcal{P}y$

	1	...	k	...
...
x	...	x	z	? ... ?
...	x	z ... z
...	? ... ?
y
...	...	?	...	x ... x
...	...	z	?	...
z	...	?	?	...
...

being decisive on $\{z, y\}$, there must exist a profile ϕ , where $z\mathcal{R}_\phi y$. Introduce y directly above or below z for all $\neq 1$ in v , so that the relative order of y and z matches that of ϕ to construct the profile χ , where the \dagger denotes a possible location for y (Fig. 18). By IIA and $\chi|_{y,z} = \phi|_{y,z}, z\mathcal{R}_\chi y$. By IIA and $\chi|_{x,z} = v|_{x,z}, x\mathcal{P}_\chi z$. Therefore, by transitivity $x\mathcal{P}_\chi y$. Now, keeping their relative ranks with respect to x fixed, arbitrarily rearrange the ranking of y, z for every agent (other than k and 1), and raise z over x in k 's ranking. This delivers the profile ψ , again the ? denotes possible locations of y (Fig. 19). By IIA and $\psi|_{x,y} = \chi|_{x,y}, x\mathcal{P}_\psi y$. By IIA and $\psi|_{x,z} = \tau|_{x,z}, z\mathcal{R}_\tau x$. Therefore, by transitivity, $z\mathcal{P}_\mu y$. This holds for every way that y can be distributed

across the question marks, which exhausts all relative orders of z and y (where $zP_k y$). By IIA, this makes k weakly decisive on (z, y) .

Since j is weakly decisive on (y, z) ; k is weakly decisive on (z, y) ; and neither j nor k are identical to agent 1 (by the construction of profiles ξ and τ), we also show that $j = k$. Suppose not. Then, there exists a profile ω , such that $yP_i z$ and $zP_j y$. By their weak decisiveness on their respective ordered pairs, we have $yP_\omega z$ and $zP_\omega y$. Contradiction, therefore, $j = k$.

Step 6. Finally, we show that $i = j$ which completes the proof. Suppose that $i \neq j$. Then, since $i \neq 1$ and $j \neq 1$, there exists a profile Γ , where $yP_i xP_i z$ and $zP_j y$. Since i is decisive on $\{x, y\}$ and $\{x, z\}$, $yP_\Gamma x$, $xP_\Gamma z$ so by transitivity $yP_\Gamma z$. Since j is decisive on $\{y, z\}$, $zP_\Gamma y$. Contradiction. Therefore, $i = j$.

Since x, y, z were chosen arbitrarily, this shows that for any triple of distinct alternatives in X , if 1 is not decisive on any pair, then some other agent is decisive on all pairs from those alternatives. □

Theorem 3 has an almost immediate corollary that any function satisfying the Arrovian conditions on a domain of the form $\{P_1\} \times \mathbf{P}^{n-1}$ must be, such that agent 1 is decisive on at least one pair of alternatives.

Corollary 3.1 *For any $n \geq 2$, $|X| \geq 3$, and $P_1 \in \mathbf{P}$, if a function $f : (\{P_1\} \times \mathbf{P}^{n-1}) \rightarrow \mathbf{R}$ satisfies IIA, WP, and ND, then agent 1 is decisive on at least one pair of alternatives from X .*

Proof Assume that $f : (\{P_1\} \times \mathbf{P}^{n-1}) \rightarrow \mathbf{R}$ satisfies IIA, and WP. Assume that agent 1 is not decisive on any pair of alternatives in X to show that f violates ND.

We prove this by induction on the set of alternatives. For the base case, we have an agent $i \neq 1$ who is decisive on all pairs from $\{a_1, a_2, a_3\}$ by Theorem 3. For the inductive step, we show that if i is decisive on all pairs from $\{a_1, \dots, a_n\}$, then i is decisive on all pairs from $\{a_1, \dots, a_n, a_{n+1}\}$. Assume that i is decisive on all pairs from $\{a_1, \dots, a_n\}$. Then, i is decisive on the pair $\{a_{n-1}, a_n\}$. Consider the triple $\{a_{n-1}, a_n, a_{n+1}\}$. By Theorem 3 and the assumption that 1 is not decisive on any pair in X , there exists an agent $j \neq 1$ who is decisive on all pairs from this triple. Therefore, j is decisive on $\{a_{n-1}, a_n\}$. However, $i = j$ for the same reason as given at the end of step 5 in the previous proof. Therefore, i is also decisive on all pairs from $\{a_{n-1}, a_n, a_{n+1}\}$ which means i is decisive on all pairs from $\{a_1, \dots, a_n, a_{n+1}\}$, since if $i \neq 1$ is decisive on a pair $\{x, y\}$ and a pair $\{y, z\}$, then i is decisive on $\{x, z\}$ (otherwise, there would exist a profile, where zRx , but also $xP_i yP_i z$ which would deliver xPy , yPz by i 's decisiveness on (x, y) and (y, z) , which by transitivity would deliver xPz ; contradiction). This suffices to establish that i is a dictator, which proves the corollary. □

3 Concluding discussion

Theorems 1 and 2 demonstrate that domains of profiles constructed by holding the preferences of one agent fixed are Arrow consistent. The functions used to demonstrate this result had the prima facie undesirable property of using the preferences of the fixed agent to determine the ‘overall’ structure of the aggregate ordering, and

allowing the agents with varying preferences to ‘fill in the gaps’. Theorem 3 showed that any function satisfying IIA, WP, and ND on such domains had to be such that if the agent with the fixed preferences does not determine the social preference of any pair of alternatives within an arbitrary triple, some other agent determines the social preference over the entire triple across the domain. Alternatively, by contraposition, for any triple of alternatives, if the social preference over them is not to be determined by a single agent across the domain, then the agent with the fixed preferences will determine the social preference over at least one pair from the triple across the domain. Thus, the influence on the social preference over pairs in a triple of alternatives is distributed between the agent with the fixed preference and some other agent. It is worth considering whether there are any instances, where such functions could be used reasonably.

It seems that they cannot in the standard context of aggregating the preferences of individuals. There is no obvious reason to privilege the agent with the fixed preference, and nor is it desirable for an arbitrary triple of alternatives to allow another agent determine the social preference on them. It is an open question as to whether a hybrid welfarist/non-welfarist account of social justice could be developed in view of the restrictions Theorem 3 highlights. However, it is not implausible that there are some cases of social choice where one might want to use some objective criterion to determine the ‘overall’ social ordering whilst still giving individuals freedom to determine the respective orderings within the ‘gaps’ in the manner of the functions defined above.

Such functions are more reasonable in the case of aggregating the ‘preferences’ of theoretical virtues over competing scientific theories to an all-things-considered ranking of them. Okasha (2011, p. 98) considers a possible ‘empiricist’ response to the problem he posed. The empiricist might claim that accuracy/fit-with-data *should* be a dictator, thereby rejecting ND in this context. I think that Okasha is right to dismiss this response on the grounds that such an approach faces the ‘problem of over-fitting’ (Forster and Sober 1994): the idea is that our data contain noise, and if we always prefer the theory that best fits the data, then we may be led to preferring radically complicated theories that have been generated precisely to match noise. Theorems 1 and 2 provide the empiricist with a solution. If the noise-matching theories are too complex, then they appear further down in the ranking provided by simplicity. Partitioning the ranking supplied by simplicity, and defining the appropriate aggregation function using this partition, will then ensure that they appear further down in the all-things-considered ranking. The empiricist is then free to claim that accuracy should determine the ranking of the remaining alternatives.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Arrow, K. (1963). *Social choice and individual values* (2nd ed.). New York: Wiley.
- Barberá, S. (1980). Pivotal voters. *Economics Letters*, 6(1), 13–16.
- Bradley, S. (2017). Constraints on rational theory choice. *The British Journal for the Philosophy of Science*, 68, 639–661.
- Breton, M. L., & Weymark, J. A. (2011). Arrovian social choice theory on economic domains. In J. Kenneth, A. S. Arrow, & K. Suzumura (Eds.), *Handbook of social choice and welfare* (pp. 191–299)., 2 Amsterdam: Elsevier.
- Cresto, E., & Tajer, D. (2019). Confirmation holism and theory choice: Arrow meets Duhem. *Mind*.
- Fishburn, P. C. (1976). Dictators on blocks: Generalizations of social choice impossibility theorems. *Journal of Combinatorial Theory, Series B*, 20(2), 153–170.
- Forster, M., & Sober, E. (1994). How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *The British Journal for the Philosophy of Science*, 45(1), 1–35.
- Gaertner, W. (2001). *Domain conditions in social choice theory*. Cambridge: Cambridge University Press.
- Gaertner, W., & Wüthrich, N. (2016). Evaluating competing theories via a common language of qualitative verdicts. *Synthese*, 193(10), 3293–3309.
- Geanakoplos, J. (2005). Three brief proofs of Arrow's impossibility theorem. *Economic Theory*, 26(1), 211–215.
- Kolm, S.-C. (1996). Rational just social choice. In Arrow, K. J., Sen, A., and Suzumura, K., editors, *Social Choice Re-Examined: Volume 2 Proceedings of the IEA Conference held at Schloss Hernstein, Berndorf, Vienna, Austria*, pp. 167–196. Palgrave Macmillan, London.
- Marcoci, A., & Nguyen, J. (2019). Objectivity, ambiguity, and theory choice. *Erkenntnis*, 84(2), 343–357.
- Morreau, M. (2014). Mr. Fit, Mr. Simplicity and Mr. Scope: From social choice to theory choice. *Erkenntnis*, 79(6), 1253–1268.
- Morreau, M. (2015). Theory choice and social choice: Kuhn vindicated. *Mind*, 124(493), 239–262.
- Okasha, S. (2011). Theory choice and social choice: Kuhn versus Arrow. *Mind*, 120(477), 83–115.
- Okasha, S. (2015). On Arrow's theorem and scientific rationality: Reply to Morreau and Stegenga. *Mind*, 124(493), 279–294.
- Rizza, D. (2014). Arrow's theorem and theory choice. *Synthese*, 191(8), 1847–1856.
- Sakai, T., & Shimoji, M. (2006). Dichotomous preferences and the possibility of arrovian social choice. *Social Choice and Welfare*, 26(3), 435–445.
- Sen, A. (1970). The impossibility of a paretian liberal. *Journal of Political Economy*, 78(1), 152–157.
- Stegenga, J. (2013). An impossibility theorem for amalgamating evidence. *Synthese*, 190(12), 2391–2411.
- Stegenga, J. (2015). Theory choice and social choice: Okasha versus Sen. *Mind*, 124(493), 263–277.
- Weymark, J. A. (2011). On Kolm's use of epistemic counterfactuals in social choice theory. In M. Fleurbaey, M. Salles, & A. J. Weymark (Eds.), *Social ethics and normative economics: Essays in honour of Serge-Christophe Kolm* (pp. 279–301). Berlin: Springer.
- Yu, N. N. (2012). A one-shot proof of Arrow's impossibility theorem. *Economic Theory*, 50(2), 523–525.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.