Public Choice (2019) 179:145–164
https://doi.org/10.1007/s11127-018-0604-7

A defense of Arrow's independence of irrelevant alternatives

John W. Patty[1] · **Elizabeth Maggie Penn**[1]

Received: 3 August 2018 / Accepted: 27 August 2018 / Published online: 7 September 2018

Abstract
Since the publication of *Social Choice and Individual Values*, Kenneth Arrow's independence of irrelevant alternatives (IIA) axiom has drawn criticism for being too strong a requirement of a collective choice rule. In this article, we detail and counter some of the criticisms. We present two axioms (one cardinal and one ordinal) that are equivalent to Arrow's IIA, discuss an implication of IIA for transitive social choice, and argue that violations of IIA do indeed constitute a perversity. We claim that violations of IIA are particularly troubling in contexts where many alternatives are considered simultaneously, good information is available about the ranking of each alternative with respect to each criterion being aggregated, final decisions can be scrutinized and revisited, and/or the correlation between how the various criteria rank alternatives is low. While a mass election is precisely a decision-making scenario that satisfies *none* of these conditions, we argue that the normative appeal of IIA is maximized for aggregation problems that can be revisited and revised, and that involve objective and varying criteria (e.g., routine administrative and judicial decisions).

Keywords Arrow's theorem · Social choice · Independence of irrelevant alternatives · IIA · Measurement

1 Introduction

Arrow's condition of *independence of irrelevant alternatives* (IIA) requires that the rule a group utilizes in order to evaluate a set of alternatives must depend only on the group members' preferences for those alternatives. Put differently, a choice rule cannot respond to members' preferences for "irrelevant" alternatives not contained in some set S when rendering a judgment over the alternatives in S. Arrow's theorem proves that for transitive decision rules, the condition of IIA is inconsistent with unrestricted domain (the

We are grateful to Donald Saari and William Shughart for their thoughtful suggestions and comments.

✉ Elizabeth Maggie Penn
elizabeth.m.penn@gmail.com; epenn@emory.edu

John W. Patty
jwpatty@gmail.com

[1] Political Science, Emory University, Atlanta, USA

assumption that all individual preferences are allowable), non-dictatorship, and the Pareto principle. Since the publication of *Social Choice and Individual Values*, that condition has been the most controversial of Arrow's axioms, in part because Arrow's original monograph contained a claimed illustration of IIA that actually was unrelated to the IIA condition.[1] While Arrow's error demonstrated that the IIA condition is nuanced and potentially difficult to conceptualize, a number of scholars have voiced other criticisms of the axiom.

In this article we provide an independent defense of IIA (Sect. 3) and respond to claims that IIA is itself at odds with collective rationality and/or restricts collective preferences from being sensitive to context.[2] In the lead up to our defense of IIA, we present several technical results concerning the IIA axiom, many of which are known or easily derived. These results are intended to bolster our substantive points concerning the IIA axiom.

1.1 Ignoring individual rationality

Donald Saari has criticized IIA as requiring the aggregation rule to ignore information about whether individuals' preferences are in fact transitive:

> IIA strictly forbids a procedure from using any (information about the transitivity of individual preferences); it must strictly concentrate on the rankings for each pair without reference to any other pairwise rankings.[3]

Accordingly, Saari is concerned in particular with how IIA undermines the assumption of voter rationality. He argues that because IIA requires information only on binary comparisons, any procedure satisfying IIA on an unrestricted preference domain can't distinguish between certain preference profiles in which voters are rational versus ones in which voters have intransitive preferences: the IIA procedure must generate the same outcome at both the rational and irrational profile. Consequently, the assumption that voter preferences are transitive is moot because the information utilized by an IIA procedure across transitive voter preference profiles must be consistent with the information used across intransitive profiles. In light of this, Saari argues that it's hardly surprising that IIA procedures generate intransitive social preferences or dictatorship.

1.2 Context independence

Amartya Sen provides a different criticism of IIA, one that takes issue with the Arrovian agenda more broadly.[4] In part, he argues that individual preference orderings are the wrong informational base for making social welfare judgments. More specifically, he presents a defense of the idea that choices may depend on external factors, so that the *context* of choice matters. For example, if a person chooses x from the set $\{x, y\}$ and y from the set $\{x, y, z\}$, many conditions regarding the internal consistency of choice are violated (including a choice functional definition of independence, as the presence of z affects the revealed

[1] Blin (1976, p. 96) and Arrow (1963, pp. 26–27). Ray (1973) disentangles numerous, commonly invoked independence axioms. See Bordes and Tideman (1991) for an extended discussion of Arrow's mistake.

[2] The literature contains other criticisms of the IIA axiom. We discuss and respond to several of these criticisms in depth in Chapter 3 of Patty and Penn (2014).

[3] Saari (1998), p. 248, Saari (1994), pp. 327–333.

[4] Sen (2014, pp. 37–39)

choice between x and y). If y represents eating the last apple in the fruit basket and x represents eating nothing, our person may want to leave the apple for someone else. However, if another apple z is added to the basket, our person may eat one of the apples (choosing y). Thus, Sen argues, the menu of options (including the presence of "irrelevant" alternatives) may indeed affect choice behavior in a natural way, with no perversity entailed.[5] Sen takes this as a charge to extend Arrow's Theorem to a setting in which no internal consistency of choice is assumed and/or, relatedly, explicitly to allow factors such as *interpersonal comparability*, *liberty*, and *rights* to affect collective choice.[6] We return to a discussion of Sarri's and Sen's cogent and insightful criticisms of IIA—and why we think they are unfounded in the contexts in which we feel that IIA is most compelling in—in Sects. 2 and 3, respectively. First, however, we note a few other defenses of the axiom offered previously.

1.2.1 Strategy-proofness

Jean-Marie Blin defends the axiom on the basis of its game-theoretic foundations, noting Mark Satterthwaite's finding that a one-to-one correspondence exists between strategy-proof procedures and social welfare functions satisfying Pareto and IIA.[7] Blin also argues that such a "game view" of social choice informed Condorcet's search for a true social preference relation in his famous jury theorem. Blin writes, "Gibbard and Satterthwaite have established what Condorcet intuitively felt: to be strategy-proof, a voting procedure must verify the independence of irrelevant alternatives. If not, then other alternatives become very relevant since they can be used for individual preference misrepresentation."[8]

1.3 Eliminating ambiguities

In response to a 1969 piece by Bengt Hansson that is critical of IIA (Hansson 1969), Peter Fishburn defends the axiom with the argument that allowing irrelevant alternatives to matter opens a Pandora's box with respect to even defining the social choice problem being considered. He writes, "[i]f in fact the social choice can depend on infeasibles present in the 'voting,' which infeasibles should be used? For with one infeasible set, feasible x might win, whereas feasible $y \neq x$ might win if some other infeasible set were used. Hence, the inclusion of infeasibles in the decision process introduces a rather horrifying ambiguity unless something like Arrow's independence condition is used" Fishburn (1970, p. 933). Fishburn goes on to argue that social transitivity is less compelling than independence,[9] particularly in light of arguments provided by Arrow in the second edition of his book. However, in later work Hansson rightly notes "not all of the arguments Arrow advances for

[5] This example is from Sen (1993), along with an extension of Arrow's Theorem that eschews the notion of social rationality.

[6] Building directly off Sen's argument concerning the effect of external factors on choice, Roberts (2009) weakens independence to allow certain irrelevant alternatives to provide information about a choice set being considered and to endogenously affect social choice. Luce and Raiffa (1957) provide a similar argument against independence in the context of choice functions, considering it plausible for the menu of choices itself to provide information about the relative ranking of alternatives.

[7] This result can be found in Satterthwaite (1975), p. 204.

[8] Blin (1976), p. 103. In addition to Satterthwaite (1975), Blin is referring to Gibbard (1973).

[9] A position we agree with for related (but different) reasons, as we discuss briefly in Sect. 3 of this article and in greater depth in Patty and Penn (2014).

the independence of irrelevant alternatives are really arguments for the same thing" (Hansson 1973, p. 41). Hansson details confusion stemming from various authors' attempts to translate independence conditions on social preference orderings to those on group choice functions (including Arrow's own attempts), and disentangles interpretations of the IIA condition in these two environments.

1.4 Relationship to other desirable axioms

Other axioms have been defended that have the flavor of Arrovian independence. An example is Sen's *minimal liberalism*, which assumes at least two individuals that are each decisive over distinct pairs of alternatives. Sen argues for the reasonableness of his liberalism axiom by claiming that certain choices are purely personal, and people should be able to exercise sole discretion over such choices. Blau (1975) notes that these pairs of alternatives are independent sets, in the sense that if preferences agree on them, then the social ranking must agree on them also. Thus, Blau argues, Sen's liberalism is in fact a weak independence condition. Moreover, Sen defends liberalism above Pareto by arguing that the pair in question may concern only the decisive individual, whereas social preferences arising from the Pareto condition may stem from "nosy" motives in which people have preferences over things that should be none of their business (for example, whether someone else chooses to read a lascivious book or not). Blau argues that Sen's independence is too strong for the claims about liberalism that he makes because constructing Sen's paradox requires an individual to be decisive over a pair that someone else feels more strongly about, thus contradicting the defense that such choices are "purely personal."[10] Finally, and relatedly, IIA is closely akin to various notions of treating all alternatives fairly (i.e., notions of *neutrality*), as discussed by Sen (1970a), Fishburn (1973), Ubeda (2003), Campbell and Kelly (2007), Gailmard et al. (2008), and Penn et al. (2011).

With the criticisms and some alternate defense of IIA detailed, we now turn to setting the stage for the more detailed rendition of our argument, beginning with some notation to help clarify the exposition when it gets into the weeds.

1.5 Formal preliminaries

Let X denote a finite set of k alternatives and N denote a finite set of $n > 1$ people. Each person $i \in N$ has preferences over the elements of X that are represented by a weak order, $\succeq_i$. We denote the asymmetric portion of $\succeq_i$ by $\succ_i$ and the symmetric portion by $\sim_i$. Let $\mathcal{R}$ denote the set of all weak orderings of X. A preference profile, ρ, maps N into $\mathcal{R}$; it assigns a weak order to each individual $i \in N$. Letting $\mathcal{B}$ denote the set of complete binary relations on X, a *preference aggregation rule* f maps the set of all possible profiles $\mathcal{R}^n$ into $\mathcal{B}$; to each preference profile it assigns a complete binary relation over the alternatives. We will term $f(\rho)$ the *social preference relation* and, for any pair of alternatives $x, y \in X$ will write $x \succeq_{f(\rho)} y$ if $f(\rho)$ ranks x at least as high as y, $x \succ_{f(\rho)} y$ if $f(\rho)$ ranks x strictly higher than y, and $x \sim_{f(\rho)} y$ if $f(\rho)$ ranks x and y the same. If $f(\rho)$ is transitive for all $\rho \in \mathcal{R}^n$, then it is a *social welfare function*.

[10] Sen (1970c). Both Blau and Gibbard present constructive modifications to Sen's axiom in which rights to decisiveness may be disallowed or waived ("martyrdom" for Blau and "alienable rights" for Gibbard). See Gibbard (1974) and Blau (1975).

1.6 Independence of irrelevant alternatives

Let S denote a nonempty subset of X: Arrow's formulation of the IIA condition argues that the elements of $X \setminus S$ should be irrelevant to the determination of social preference on S. Let $\rho|_S$ be *the restriction of* ρ *to* S, so that $\rho|_S$ describes individuals' preferences only over the alternatives in S. Similarly, let $f(\rho)|_S$ be the restriction of the social preference relation to S, so that it describes societal preferences only over the elements of S.

Following Blau's terminology, we will say that if for some $S \subseteq X$ we have for *any* pair of profiles ρ, ρ' that $\rho|_S = \rho'|_S$ implies $f(\rho)|_S = f(\rho')|_S$, then the subset S is *independent*. Arrow's IIA is then defined as follows:

Definition 1 A preference aggregation rule f satisfies **independence of irrelevant alternatives** if *every* set $S \subseteq X$ is independent.

Independence of irrelevant alternatives often is termed *binary independence*, because, as May (1954) noted, the condition can be rewritten solely in terms of independence of *pairs* of alternatives, as opposed to all subsets of alternatives. Thus, the above definition of IIA (Definition 1) is equivalent to the following definition.

Definition 2 A preference aggregation rule f satisfies **independence of irrelevant alternatives (IIA)** if *every* pair $\{x, y\} \in X \times X$ is independent.

The equivalence of Definitions 1 and 2 stems from the fact that if two profiles are identical on some subset S, then they also are identical on all pairs of elements in S. Therefore, if independence on pairs holds, then independence on any larger subset likewise must hold. To be clear, because the relationship plays a subtle but important role in our defense of IIA below, May's useful and powerful definition (Definition 2) is equivalent to Arrow's IIA, but its simplicity can overstate the leap one *must* take in imposing Arrow's axiom. That is, Arrow's IIA implies May's definition, but one can start with an apparently much weaker (or, perhaps, more palatable) axiom and arrive at the same conclusions. Formally or logically, this is a distinction without a difference, but as this article manifests, the arguments about axioms are not solely on logical grounds—they often are about reasonableness, an arguably more important and interesting concept. With this in hand, we turn first to Saari's compelling argument against the reasonableness of Arrow's IIA.

2 IIA, transitivity and information loss

Saari's objection to IIA(Saari 2001, 2006) essentially is that it forces the aggregation rule to ignore the structure (or "information") contained within the individual criteria. By requiring the collective ordering to respond only to how the criteria rank *pairs* of alternatives, IIA rules out aggregation rules (such as Borda count) that leverage the transitivity of the individual criteria reliably to return another transitive ordering.

This is a provocative and intriguing objection. We respond to it in four ways, each of which we personally find persuasive, but even the combination of which we do not assert is dispositive: as we return to in the conclusion of the article, we agree with Saari that many meaningful situations in which Arrow's IIA is unacceptably restrictive certainly can be found.

Our response begins with a consideration of aggregation rules that generate a unique "best alternative" and show (Theorem 1) that, in such settings, IIA is in fact *sufficient* for collective transitivity.[11] Second, we present and discuss one of the analogues to IIA in the choice functional setting that does not rely on the "rationality" of the choice function in question. We believe that, viewed in conjunction, these responses clarify that the challenge posed by IIA is not the consequence of pursuing a transitive social ordering.

We then discuss a strictly weaker notion of independence proposed by Blau (1971) that requires independence only when considering subsets of size k; this independence condition does not logically require binary independence, unless $k = 2$ is chosen. Blau shows, however, that this weaker independence implies binary independence. Thus, independence on any strict subset of alternatives is equivalent to independence on pairs.

Finally, and as we transition to responding to Sen's objections, we discuss a cardinal version of IIA. When preferences are cardinal, then not only are they by definition transitive, but the preference aggregation rule can respond to information that is sufficient to recover the entire transitive ordering of any criterion. We argue that the two responses imply that IIA, per se, does not render the individual criteria essentially intransitive. With this road map described, we now move to discuss our four responses to Saari's argument in turn.

2.1 IIA and transitivity

Our first response to Saari's objection to IIA focuses on preference aggregation rules that return an unambiguous choice in the sense that they always return a unique alternative ranked above all others—we refer to such a preference aggregation rule as satisfying *topness*.

Definition 3 A preference aggregation rule f satisfies **topness** if, for all ρ, there exists some $x_\rho \in X$ such that $x_\rho \succ_{f(\rho)} y$ for all $y \in X$.

Note that a preference aggregation rule f that satisfies topness need not be transitive; such a function might return a binary relation that contains one or more cycles "under" the top-ranked alternative. For example, a preference aggregation rule that, for some ρ and $X = \{x, y, z, w\}$, returns $x \succ_{f(\rho)} y$, $x \succ_{f(\rho)} z$, $x \succ_{f(\rho)} w$, $y \succ_{f(\rho)} z$, $z \succ_{f(\rho)} w$, $w \succ_{f(\rho)} y$, satisfies topness in spite of being cyclic. In contrast to this possibility, the next axiom describes a preference aggregation rule that never returns either cycles or indifference.

Definition 4 A preference aggregation rule f satisfies **strict ordering** if, for all ρ, $f(\rho)$ is a strict order, or a complete, antisymmetric, and transitive binary relation, on X.

Finally, we utilize the following axiom akin to citizens' sovereignty that we term "top feasibility." It states that for any alternative, there exists some profile that ranks it uniquely "top." Note that this axiom is implied by, but weaker than, Arrow's axiom of weak Pareto, which we do not assume.

[11] Our result also utilizes a citizens' sovereignty-type condition.

Definition 5 A preference aggregation rule *f* satisfies **top feasibility** if, for any alternative $x \in X$, there exists a profile ρ with $x \succ_{f(\rho)} y$ for all $y \in X$.

With these axioms in hand, the following theorem (which is straightforward but, to the best of our knowledge, not contained in the literature, at least as stated) states that IIA, topness, and top feasibility imply strict ordering.[12]

Theorem 1 *With three or more alternatives, any preference aggregation rule f satisfying IIA, topness, and top feasibility must also satisfy strict ordering.*

Proof Let *f* be a preference aggregation rule satisfying IIA, topness, and top feasibility. We will show that *f* satisfies strict ordering directly by showing that $f(\rho)$ is a strict order for any arbitrary preference profile $\rho \in \mathcal{R}^n$. Accordingly, fix a preference profile $\rho \in \mathcal{R}^n$.

We first show that $f(\rho)$ does not admit any indifference. Suppose by way of contradiction that for some pair x, z and some profile ρ we have $x \sim_{f(\rho)} z$. By top feasibility there exists another profile, ρ^z with z ranked top. Now construct a third profile, ρ' such that for all $y \in X \setminus \{x, z\}$, $\rho'|_{yz} = \rho^z|_{yz}$ and $\rho'|_{xz} = \rho|_{xz}$. By IIA, $z \succ_{f(\rho')} y$ for all $y \in X \setminus \{x, z\}$ and $z \sim_{f(\rho')} x$. However, this contradicts topness. Thus, *f* admits no social indifference.

We now show that $f(\rho)$ is a strict order. Suppose not, so that there exists a triple $\{x, y, z\}$ with $x \succ_{f(\rho)} y$, and $y \succ_{f(\rho)} z$ and $z \succ_{f(\rho)} x$ (as *f* admits no indifference, a violation of transitivity implies such a cycle). By topness, there must exist a distinct fourth alternative q_1 with q_1 top at profile ρ. Note that if $k = 3$ then our proof is complete at this point, as the cycle violates topness.

Now reconsider profile ρ^z which ranks z top, and which we know exists by top feasibility. Construct a third profile, ρ^1 such that $\rho^1|_{xyz} = \rho|_{xyz}$ and $\rho^1|_{q^1z} = \rho^z|_{q^1z}$. By IIA, $z \succ_{f(\rho^1)} q^1$, and $x \succ_{f(\rho^1)} y$, and $y \succ_{f(\rho^1)} z$, and $z \succ_{f(\rho^1)} x$. By topness, there exists a $q^2 \notin \{x, y, z, q^1\}$ with q^2 ranked top at $f(\rho^1)$. Again, if $k = 4$ the proof is complete.

The proof proceeds by exhausting the remaining set of alternatives to contradict topness. We'll proceed with one more step to demonstrate. Construct a fourth profile ρ^2 with $\rho^2|_{xyzq^1} = \rho^1|_{xyzq^1}$ and $\rho^2|_{q^2z} = \rho^z|_{q^2z}$. By IIA, $z \succ_{f(\rho^2)} q^2$, and $x \succ_{f(\rho^2)} y$, and $y \succ_{f(\rho^2)} z$, and $z \succ_{f(\rho^2)} x$, and $z \succ_{f(\rho^2)} q^1$. By a similar argument, there exists a $q^3 \notin \{x, y, z, q^1, q^2\}$ with q^3 ranked top at $f(\rho^2)$. The proof concludes when we reach q^{k-3}, where k is the total number of alternatives. At this point we violate topness because there exists no possible q^{k-2} at the top of $f(\rho^{k-3})$. This contradicts our assumed violation of transitivity. Thus, $f(\rho)$ is a strict order for all profiles ρ. □

Theorem 1 offers one rejoinder to Saari's objection to IIA. If an IIA aggregation method is strongly consistent with "revealed collective choice" in the sense that the method always returns a unique top-ranked alternative, then the method also necessarily strictly orders all of the alternatives. Thus, IIA is not inconsistent with collective rationality. Indeed, emboldened by topness, IIA *guarantees* that social preferences will be rational and strict. Furthermore, our proof did not utilize any assumption of individual rationality. While IIA loses

[12] Of course, the converse is not true. Note that Theorem 1 is independent of choice theoretic results such as the equivalence between the weak axiom of revealed preference (WARP) and rationalizability of choice (Arrow 1959), because the choice set is held fixed in Theorem 1 (and similarly for conditions α and γ in Sen 1971). Strict ordering of *f* implies that WARP holds at every profile ρ, but satisfaction of WARP does not imply that *f* satisfies IIA, topness, top feasibility, or strict ordering.

(some) information about individual rationality—and as we discuss below, this loss helps drive Arrow's result—IIA *secures* collective rationality despite the loss of this information.

Saari argues that IIA renders the search for collective rationality hopeless, but Theorem 1 shows that IIA along with topness—essentially a weak tie-breaking assumption—*implies* collective rationality. Put another way, any aggregation method that always generates a unique choice that is *not* collectively rational over the entire set of alternatives *must* violate IIA. Furthermore, because IIA and topness imply transitivity, any procedure that satisfies these axioms is (by Wilson's theorem) normatively undesirable.[13] Our interpretation of the problems stemming from IIA is that it resolves the tension between collective rationality and democracy in an "all or nothing" way. If an IIA system is not required to produce a unique top then we are left with potential cycles (e.g., majority rule) or ties (e.g., unanimity rule) at the top of the social preference ordering. Any resolution of such indeterminacy must favor one person over another. IIA takes this favoritism to the extreme by forcing consistency in how ties are broken across preference profiles, which loses information and leaves us with dictatorship.

We feel Theorem 1 buttresses Sen's claim that "...the real source of the impossibility problem is the tension between the informational eschewal implicitly imposed by Arrow's set of axioms and the demands of discriminating social choice also entailed by the same axioms (Sen 1993, p. 514, fn 38)". In this respect, our theorem is simply another instantiation of Arrow's theorem in an environment in which internal consistency of choice is not imposed. Sen (1993) uses the term "internal consistency of choice" to refer to the requirement of societal rationality in the form of a transitive social preference ordering or societal choices satisfying the weak axiom of revealed preference. Both Sen (1993) and Denicolò (1985) prove versions of Arrow's Theorem for choice correspondences; their theorems eschew internal consistency because they place no explicit rationality conditions on the choice correspondences considered, such as requiring them to satisfy the weak axiom of revealed preference (or, equivalently, be transitively rationalizable).[14] As our second response to Saari's critique, we now briefly describe Denicolò's important result.

2.2 Independence and choice without rationality

Sen argues that when we move from the setting of generating a collective preference relation to one of simply generating a collective choice, the question of whether a choice rule respects collective rationality becomes irrelevant when the set of alternatives is fixed (Sen 1993). To represent collective choice, and letting $\mathcal{X}$ denote all nonempty subsets of X, a *social choice correspondence* is a correspondence $C : \mathcal{R}^n \rightarrow \mathcal{X}$. In this setting, Denicolò defines three axioms: *Pareto optimality*, *dictatorial*, and *independence*. The Pareto optimality and dictatorial axioms are fairly standard and defined below.

[13] Wilson (1972) proves that IIA and a weak citizens' sovereignty condition (that is implied by topness) imply a social welfare function that is either dictatorial, inverse dictatorial (reversing an individual's preferences, which Wilson also terms "dictatorial"), or null (generating a tie over all alternatives, which is excluded here as it would violate topness).

[14] It is useful to note that Arrow himself defined IIA in choice-functional terms. Specifically, if for some $S \subseteq X$ individuals' preferences over S are unchanged, then the choice set $C(S)$ must remain unchanged (Arrow 1963, p. 27; Sen 1993, p. 510).

Axiom 1 A social choice correspondence C is **Pareto optimal** if, for all $x \in X$ and any $\rho \in \mathcal{R}^n$,

$$x \succ_i y \forall i \in N \Rightarrow y \notin C(\rho).$$

Axiom 2 A social choice correspondence C is **dictatorial** if there exists $d \in N$ such that

$$x \succ_d y \Rightarrow y \notin C(\rho).$$

With these in hand, the independence axiom is the axiom of interest to us in this section.

Axiom 3 A social choice correspondence C is **independent** if, for all $y, z \in X$ and all $\rho, \rho' \in \mathcal{R}^n$ with $\rho|_{yz} = \rho'|_{yz}$,

$$x \in C(\rho) \text{ and } y \notin C(\rho) \Rightarrow y \notin C(\rho').$$

The main result of Denicolò (1985) is that independence and Pareto optimality are satisfied only by dictatorial social choice correspondences.

Theorem 2 (Denicolò 1985) *If a social choice correspondence C is Pareto optimal and independent, then C is dictatorial.*

Because they both focus on choice, allowing for both single-valued and multi-valued choice, we argue that Theorem 1 and Denicolò's Theorem jointly justify Sen's conclusion regarding the relative unimportance of transitivity in Arrow's theorem:

> Neither internal consistency of social choice, nor any restricted structure of 'social preference' entailed by alleged 'social rationality,' is the source of the impossibility problem identified by Arrow. (Sen 1993, p. 514)

Thus, while we agree with Saari's assertion that Arrow's IIA requires that the preference aggregation rule ignore at least some of the information contained in the transitivity of the individual criteria, we argue that the effect of this information loss is not necessarily the loss of collective rationality. If these results are not persuasive, we now turn to consider the relationship between May's clear and strong "binary" definition of IIA and the degree to which its clarity and strength are (not) required to obtain Arrow's result.

2.3 IIA and binariness

Saari is not the first to object to IIA as essentially requiring that the preference aggregation rule respond only to information about *pairs* of alternatives and, accordingly, requiring it to lose information about how the members of those pairs fare with respect to other alternatives. For example, Gerry Mackie states that "[t]he IIA [condition] would better be named the pairwise comparison condition, as it requires that choices among several alternatives be carried out only with information about choices between pairs" (Mackie 2003, p. 139) and, indeed, even Arrow writes, "[k]nowing the social choices made in pairwise comparisons in turn determines the entire social ordering and therewith the social choice function $C(S)$ for all possible environments." (Arrow 1951, p. 28) While this objection to, and characterization of, IIA is true, we feel that it is misguided, *because it makes it appear that allowing the rule to respond to a richer set of information would eliminate some of the perversities stemming from IIA*. In fact, these problems arise even when we allow our rules to be responsive to far more information than simple pairwise comparisons.

It was speculated by Murakami (1968) and later proved by Blau (1971), that the requirement that independence hold for all subsets of X could be weakened considerably to requiring only that independence hold on subsets of $k-1$ alternatives (i.e., subsets containing all but one of the alternatives). Thus, when there are four or more alternatives (i.e., $k \geq 4$), there is nothing explicitly about independence on *pairs* that drives Arrow's result. Blau's is an "all or nothing" type of result: independence on anything but the entire set of alternatives implies independence on pairs.

While binary independence says that the ranking of each pair must lose information about all other alternatives—a potentially large loss of information—Blau's result tells us that this same loss of information must occur even if we only require our rule to lose information about a single alternative. Phrased differently, *any* loss of information is equivalent to binary independence.

To drive home this point, we have shown in Patty and Penn (2014) that the full force of IIA can be generated by requiring a form of independence that is even less demanding than that required by Blau. Specifically, all we require is that if two profiles are identical, save for a single individual changing the ranking of two alternatives that are adjacent in his or her preference ordering, then only the social ranking of those two alternatives can change. From an "informational loss" perspective, this requirement says that the social rankings for all pairs in $X \times X \setminus \{x, y\}$ can't be affected by a single person rearranging only their ranking of x and y (thus, the information regarding this rearrangement is lost on other rankings).

Formally, a pair of profiles $\rho^1 = (\succ_1^1, \ldots, \succ_n^1)$ and $\rho^2 = (\succ_1^2, \ldots, \succ_n^2)$ is said to represent a *unilateral single-alternative deviation by individual i with respect to alternatives x and y* if there is exactly one individual i such that

1. The profiles ρ^1 and ρ^2 are identical for all individuals j other than i (i.e., $\succ_j^1 = \succ_j^2$ for all $j \neq i$), and
2. Individual i "flips" the ranking of x and y in the two profiles (i.e., $y \succ_i^1 x$ and $x \succ_i^2 y$), but $\succ_i^1$ and $\succ_i^2$ are otherwise identical.

Axiom 4 A preference aggregation rule f satisfies unilateral flip independence if, for any individual i, any pair of alternatives, x and y, and any pair of profiles ρ_1 and ρ_2 such that ρ_1 and ρ_2 are a unilateral single-alternative deviation by individual i with respect to alternatives x and y, $f(\rho_1)$ and $f(\rho_2)$ differ at most in the collective ranking of x and y: for every w distinct from x and y and every $z \in X$ it must be the case that $wf(\rho^1)z \Leftrightarrow wf(\rho^2)z$ and that $zf(\rho^1)w \Leftrightarrow zf(\rho^2)w$.

The following result implies that satisfaction of unilateral flip independence is both necessary and sufficient for IIA to hold.[15]

Theorem 3 (Patty and Penn 2014) *If X contains at least three alternatives, an aggregation rule f satisfies unilateral flip independence if and only if f satisfies IIA.*

[15] Note that we have defined and proved this result assuming strict individual preferences, although the assumption is not necessary. We follow a proof strategy similar to Geanokoplos's (2005) constructive proof of Arrow's Theorem using a pivotal voter.

Theorem 3 demonstrates a fundamental basis of the normative appeal of IIA as a choice axiom: the full strength of the axiom can be obtained by merely requiring that no single individual be able to flip two *adjacent* alternatives in his or her preference ranking and alter the social ranking of any different pair of alternatives whose relative rankings remain unchanged as a result of the flip.

We now move to consider the degree to which allowing for the criteria to be cardinal can help us circumvent IIA and/or escape Arrow's theorem.

2.4 Cardinality, comparability and measurability

Mackie (Mackie 2003, p. 134) expresses a commonly held belief when he writes that "the [IIA] condition enforces the ban on information other than individual orderings". However, as demonstrated by Sen (1970a) and others,[16] the claim that IIA "enforces the ban" on cardinal information about the criteria is untrue. Specifically, Arrow's theorem can be readily extended to allow for cardinal inputs. This extension allows a rule to treat an individual who assigns cardinal scores of $x = 3$, $y = 10$, and $z = 0$ to alternatives x, y, z at one profile as different than that same individual assigning scores $x = 15$, $y = 16$, and $z = -5$ at a different profile. The individual has the same ordinal ranking of the alternatives at both profiles ($y \succ_i x \succ_i z$), but different cardinal rankings.

We note here that this extension concerning cardinal independence holds only *so long as these inputs are not assumed to be directly comparable across individuals*. Non-comparability means that if individual i assigns $x = 5$ and individual j assigns $x = 10$ we cannot conclude that i values x more than does j. Thus, the theorem truly rests on non-comparability of the inputs. The cardinal IIA we consider essentially rules out systems such as the Borda count that treat the criteria "as if" they are directly comparable. To see this, note that if we allow the preference aggregation rule to treat the cardinal inputs as if they *are* directly comparable, then the analogue of IIA can be satisfied by many transitive, non-dictatorial, Pareto efficient preference aggregation rules: for example, a rule that assigns each alternative a score equal to the sum of individuals' utility measures and then ranks alternatives according to those scores. Treating the inputs as if they are directly comparable means that the preference aggregation rule "normalizes" individuals' cardinal preferences.[17] This is because once the individuals' "units of utility" are placed on a common scale, preferences arguably are comparable across individuals.

In this section, we prove an extension of Sen's result that independence on cardinal inputs implies independence on ordinal inputs. Thus, while allowing a rule to respond differently to different cardinal inputs appears to be a weakening of ordinal independence, it is in fact not. Our extension is to a setting in which only m-ary (with $m < k$), and not binary, independence is required, and is an amalgamation of results by Sen (1970a) and Blau (1971).[18] Our independence requirement implies that if two cardinal profiles agree in their evaluations of m alternatives, then changing the cardinal evaluation of an $m + 1$th alternative should not affect the social ranking of the original m alternatives.

[16] For example, Kalai and Schmeidler (1977) and Hylland (1980).

[17] Sen (1970a), pp. 91–98 is particularly useful with respect to this argument.

[18] After arriving at deriving this conclusion, we realized that Kalai and Schmeidler (1977) prove a similar result in their Lemma 1 on their path to proving a version of Arrow for cardinal social welfare functions.

In a cardinal preference setting, each individual is represented by *utility function*, $U_i : X \to \mathbb{R}$. Let $\mathcal{U}$ be the set of all utility functions and let $U \in \mathcal{U}^n$ denote an arbitrary *utility profile*. Any utility function U_i generates a related set of utility functions, $L_i(U_i)$, which we shall denote simply by L_i when there is no risk of confusion. The set L_i consists of the set of utility functions that we consider to be equivalent utilities under criterion i. Thus, preferences are *measurable* only up to the set L_i. The following two definitions of measurability will be used throughout (with the corresponding sets L_i being referred to as L_i^o and L_i^c).

Definition 6 Preferences are **ordinally measurable** if every element of L_i^o is a positive monotonic transformation of any other and, moreover, *every* positive monotonic transformation of an element of L_i^o is contained in L_i^o.

For example, if there are three alternatives and criterion i assigns the following utilities to them

$$U_i = (7, 0.2, 100) \in L_i^o,$$

then it must be the case that

$$U_i' = (51, 50, 52) \in L_i^o$$

must also be an element of L_i^o, as both utility functions maintain the same ordering over alternatives.

Definition 7 Preferences are **cardinally measurable and noncomparable** if every element of L_i^c is a positive affine transformation of every other element, and every positive affine transformation of any element of L_i^c is an element of L_i^c.

A positive affine transformation of a utility function is a transformation of the form

$$U_i' = a_i + b_i * U_i,$$

with $a_i, b_i \in \mathbb{R}$ and $b_i > 0$. Thus, if

$$U_i = (7, 0.2, 100) \in L_i^c,$$

then it must be the case that

$$U_i' = (701, 21, 10001) \in L_i^c \text{ (corresponding to} a = 1 \text{ and } b = 100\text{), and}$$
$$U_i'' = (35.2, 1.2, 500.2) \in L_i^c \text{ (corresponding to } a = .2 \text{ and } b = 5\text{).}$$

The “noncomparability” in this definition stems from the fact that *every* positive affine transformation of some $U_i \in L_i^c$ is in L_i^c, for all individuals i. In other words, if we transform the utility functions of different individuals using different affine transformations, the resulting set of individuals’ preferences is assumed to be an identical measure of the set of individuals’ preferences prior to the transformations.

2.4.1 Aggregation functionals

Let $\mathcal{B}_X$ be the set of all complete and reflexive binary relations on X. An *aggregation functional*, $f : \mathcal{U}^n \to \mathcal{B}_X$, maps utility profiles into an ordinal comparison of all pairs of alternatives.

Let $L = \prod_{i=1}^{n} L_i$, with L_i defined above. L is the set of all utility profiles that our rule deems as equivalent measures of individual preference, which implies that $U, U' \in L \Rightarrow f(U) = f(U')$.

2.4.2 Subsets of alternatives

For any $S \subseteq X$, let $U_i|_S$ be the utility function U_i restricted to the set S and let $U|_S$ be a utility profile restricted to S. We can then define $L_i|_S$ as the set of utility functions restricted to S that are equivalent to U_i as measures of preference *on the set* S and $L|_S$ as the Cartesian product of these sets $L_i|_S$. We also can define $f(U)|_S$ as the restriction of our aggregation functional to S.

With these preliminaries in hand, we are now in a position to state the independence axiom.

Axiom 5 An aggregation functional f satisfies ***m*-ary cardinal independence** if for any subset $S \subset X$ of size $m < k$, it is the case that if for all individuals i, $U_i|_S \in L_i|_S$ and $U'_i|_S \in L_i|_S$ then $f(U)|_S = f(U')|_S$.

Arrow's definition of binary independence (independence of irrelevant alternatives or IIA) is a case of 2-ary (binary) independence when preferences are ordinally measurable, or $L_i = L_i^o$. We are interested in the relationship between Arrow's IIA and what we term m-ary cardinal independence, which is m-ary independence when preferences are cardinally measurable and noncomparable, or $L_i = L_i^c$.

Theorem 4 *For any $m < k$ it must be the case that m-ary cardinal independence implies Arrovian binary independence.*

Proof Sen (1970b) has proven that cardinal binary implies ordinal binary. We start by showing that cardinal ternary implies ordinal binary. We then show that cardinal m-ary independence implies cardinal $(m-1)$-ary independence.

Step 1 Cardinal 3-ary (ternary) independence implies ordinal binary independence. Suppose that f satisfies cardinal ternary independence and that $k \geq 4$. Consider any two utility profiles such that for some pair x, y we have

$$U|_{\{x,y\}}, U'|_{\{x,y\}} \in L^o|_{\{x,y\}}.$$

Thus, utility profiles U, U' are ordinally identical on the set $\{x, y\}$. They likely are not cardinally identical on $\{x, y\}$ and in any case, even if they were, we know only that our aggregation functional satisfies cardinal ternary independence, not cardinal binary independence.

Now constrict a third utility profile U'' that is the following affine transformation of U': For each individual i, let

$$U''_i = a_i * U'_i + b_i$$

where[19]

$$a_i = \frac{U_i(x) - U_i(y)}{U_i'(x) - U_i'(y)} \quad b_i = \frac{U_i(y)U_i'(x) - U_i(x)U_i'(y)}{U_i'(x) - U_i'(y)}.$$

By our assumption of cardinal measurability, it must be that $f(U') = f(U'')$. Given the construction of our affine transformation, we also have that $U''|_{\{x,y\}} = U|_{\{x,y\}}$.

By the assumption that $k \geq 4$ we know that there exists an $a, b \in X$ distinct from x, y. Finally, construct a profile U^* such that $U^*(a) = U(a)$, $U^*(b) = U'(b)$, and $U^*|_{\{x,y\}} = U|_{\{x,y\}} = U'|_{\{x,y\}}$ (which we can do by the construction of our affine transformation).

By cardinal ternary independence we have

$$f(U^*)|_{\{a,x,y\}} = f(U)|_{\{a,x,y\}}$$

and

$$f(U^*)|_{\{b,x,y\}} = f(U'')|_{\{b,x,y\}}.$$

By cardinal measurability we have

$$f(U') = f(U'').$$

It follows that $f(U)|_{\{x,y\}} = f(U')|_{\{x,y\}}$, or that f satisfies Arrovian binary independence.

Step 2 Cardinal m-ary implies cardinal $(m-1)$-ary for any $m < k$. Suppose f satisfies cardinal m-ary independence and consider two utility profiles, U, U' for which there is a set S with $|S| = m - 1$ with $U|_S = U'|_S$. By the assumption that $k > m$ we know that there are at least two alternatives contained in X that are not contained in S. Call them a, b.

Construct a new profile U^* with $U^*|_{S\cup\{a\}} = U|_{S\cup\{a\}}$ and $U^*|_{S\cup\{b\}} = U'|_{S\cup\{b\}}$. By the assumption of m-ary independence we have $f(U^*)|_{S\cup\{a\}} = f(U)|_{S\cup\{a\}}$ and $f(U^*)|_{S\cup\{b\}} = f(U')|_{S\cup\{b\}}$. Thus, $f(U)|_S = f(U'|_S)$, and so f satisfies cardinal $(m-1)$-ary independence. □

Like Theorems 3 and 4 is another "all or nothing" -type result. If cardinal m–ary independence is violated, thenby merely altering the cardinal evaluation of one alternative, the social ranking of the remaining $m - 1$ alternatives can be changed. While this would appear to be a weak, easy to satisfy, condition, we show that it is in fact equivalent to (ordinal) IIA. Accordingly, *cardinality without comparability is not an escape route from the implications of binary independence or, hence, those of IIA*.

We now turn to a more full-throated and, we think, more constructive defense of IIA. Specifically, we believe that IIA is not the right target for those attempting to attack and reshape Arrow's Theorem. Rather, as a procedural/rationality quality, IIA is more normatively appealing than "collective rationality" (*i.e.*, transitivity). To conclude the article, we turn to a brief renewal and recounting of this argument, a more thorough presentation of which can be found in Patty and Penn (2014).

[19] We know that $a_i \geq 0$ for all i because U and U' order x, y the same way for any i.

3 Defending IIA: beyond preferences and elections

Our results to this point have shown that IIA, a demanding condition, follows from a number of conditions that would appear to be far weaker. These conditions allow the aggregation rule to respond to cardinal differences, and relax the assumption of binary independence. We have also shown that IIA is not incompatible with transitivity, but in fact implies transitivity (and by extension, dictator) when we require collective preference to be discriminating. At the same time (and as the title of this article suggests) we defend IIA as a sensible axiom for many decision making scenarios—scenarios wherein a single alternative must be chosen. How can this be, given our argument that independence with democracy implies indeterminacy and independence with determinacy implies dictator?

Many of the debates about IIA have focused on how collective choice should respond to changes in individual preferences.[20] More generally, Arrow's Theorem has been discussed most frequently within the context of electoral system design and what we can expect from elections in terms of social welfare and accountability (e.g., Riker 1980; Mackie 2003; McGann 2006). We believe that the appeal of the axiom is best understood by stepping back from elections and preferences. The axiom itself is not tied to either application: rather, it is an abstract logical requirement for information aggregation.

3.1 Why IIA is not compelling in elections

Our aim here is to present a defense of IIA that focuses on its relevance to multicriterion decision making more generally, rather than focusing on the narrower problem of aggregating individual preferences into a collective preference. By broadening the scope of discussion, notions such as liberty, manipulation and individual rights are not relevant; the inputs to our aggregation problem may simply be ordinal or cardinal data that are observable and verifiable.[21]

Within the more general problem of aggregating measurable criteria to make a decision and/or generate a ranking of the measured alternatives, we argue that violations of IIA indeed constitute a perversity in the aggregation of *data*. From a governance perspective—situations in which the aggregation is used to make decisions that affect multiple people—that perversity may be particularly troubling in procedures that are used frequently to evaluate many alternatives for which information about each alternative on multiple criteria can be observed. Violations of IIA are especially problematic when information about how the alternatives perform under the various relevant criteria is observable and measurable. When the criteria being evaluated cannot be observed, such as in an election wherein the "criteria" used by an aggregation procedure are individuals' preference orderings, potential violations of independence are difficult, if not impossible, to detect. Detection would involve knowing both the profile of individual preferences and a counterfactual preference profile to which our aggregation rule is applied.

[20] Changing at least one individual's preferences is required to detect a violation of Arrow's IIA axiom. Analogous independence conditions can be applied to a single preference profile (e.g., Feldman and Serrano 2008). For reasons of space, we do not discuss that line of defense for IIA in this article.

[21] In Patty and Penn (2015a, b, 2018) we offer more general discussions of social choice theory and its implications for the construction of rankings and indices (e.g., the US News and World Report rankings of colleges), and data analysis more generally.

Moreover, constructing this “counterfactual” profile is purely an intellectual exercise. If a system is reasonable and understandable to the population of voters using it, it is unlikely voters would be particularly troubled by a hypothetical violation of independence. As a case in point, the alternative vote is well known and popular even though it violates a quite weak monotonicity condition (arguably a more compelling axiom than independence).

3.2 IIA, measurement and data aggregation

As we discuss in much greater detail in Patty and Penn (2014),[22] we believe that the normative appeal of IIA is maximized for aggregation problems that occur frequently and are based on attempts to use multiple criteria to evaluate multiple possible choices. In that case, instead of aggregating voter preferences to generate an outcome, we conceive of the criteria being aggregated as measurable and known entities: for example, we might wish to aggregate student performance scores in order to generate a ranking of teacher performance. To pin down our example, each student is analogous to a voter, and the set of teachers is analogous to the set of possible voting outcomes. A student “ranks” each of their teachers according to their own performance in each teacher’s class.

Consider violations of IIA (as defined in Definition 1) within this type of measurement framework. A violation of IIA occurs when some set of alternatives, $A \subset X$, for which two sets of criteria, ρ and ρ', evaluate all alternatives in A *identically*, and yet the rankings of the alternatives, $f(\rho)|_A$ and $f(\rho')|_A$, differ. From an empirical standpoint, *a violation of IIA implies that the measurement of the alternatives in A is not identified without information about alternatives outside of A*. To make the point as clearly as possible, if f violates IIA, then situations exist in which providing evaluative information about the relevant subset of alternatives is insufficient to return an evaluation of those alternatives.

This is problematic for a number of reasons. From a practical standpoint, evaluative information may be unavailable for all of the alternatives. When such information is gathered, it changes the relative ranking of other alternatives whose measures were unchanged. To return to our student/teacher example, once Teacher A is added to our dataset, it changes the relative ranking of Teachers B through Z. Or, it might be the case that the measure of a criterion was incorrect, and must be changed. Perhaps Student i performed better in Teacher A’s class than previously thought. Again, a violation of IIA implies that this change can alter the relative rankings of the remaining alternatives, whose measures were correct.

As we discussed in our section on cardinality and our Theorem 4, those problems can be bypassed if the measures of the criteria are comparable; if, for example, a performance boost of 5 points by Student i is comparable to a boost of 5 points by Student j. However, as we and others have noted, normalizing data to be comparable across criteria involves its own set of choices and paradoxes: different reasonable methods of generating comparability across criteria will necessarily generate different rankings of outcomes.[23] One might argue that comparability of performance scores across students is reasonable, as the scores capture the same phenomenon for each student-teacher interaction.

At the same time, if the criteria being aggregated are not conceptually similar—for example, as in the comparison of congressional districting plans on the basis of each plan’s

[22] Specifically, see Chapter 3, pp. 48–63.

[23] Again, we discuss this point in greater depth in Chapter 3 of our book, Patty and Penn (2014).

adherence to criteria such as *population equality* and *compactness*—the assumption of unit comparability across criteria is less defensible.[24] There is no reason to expect measures of these criteria to be comparable or correlated, as they are practically and conceptually distinct from each other. Put simply, these criteria do not have common (much less unambiguously/uncontroversially common) scales.

As we discuss in Patty and Penn (2015a), situations with such comparability are common and Saari's insightful objections to IIA are especially relevant in such situations. When the criteria can be compared with one another, then Arrow's version of IIA is too strong and, as Saari correctly points out, unnecessarily requires that the procedure ignore information within the underlying criteria. Put another way, as cardinal criteria become more directly comparable, the appropriate definition of IIA becomes weaker, eventually becoming satisfied tautologically by any function when the criteria are considered so comparable that one simply can add them together.

3.3 IIA and political choice

The notion that IIA-type violations might be troubling and delegitimizing for the procedures vulnerable to them has historical precedent. In 1907, the new state of Oklahoma was granted five congressional seats, and the size of the House of Representatives increased by five seats to accommodate that change. After some (what were at the time) tedious calculations, it was soon discovered that, if Hamilton's method of apportionment were to be used, New York would lose a seat while Maine would gain one, even though the population numbers used in the apportionment formula were unchanged.

This phenomenon is referred to as the "New States Paradox," and vulnerability to this paradox is one (of several) arguments against Hamilton's method of apportionment, which was utilized for nearly half a century (Balinski and Young 2001, pp. 43–44) Like congressional districting, apportionment is a good example of the type of aggregation problem for which independence is compelling. They each present a problem wherein the data involved and the procedure's responsiveness to any revision in data is transparent, and both the data (e.g., the number of seats) and procedure (e.g., the apportionment method, which ultimately was changed in this case) can be revised relatively simply. With this frailty of procedure in mind, we conclude with our proposed resolution of the conundrum: forego transitivity, and require instead justifications for the chosen alternative.

3.4 One resolution

In Theorem 1, we showed that satisfaction of IIA, along with the requirement that a procedure be discriminating in the sense of picking a unique "top," implies transitivity and thus dictatorship. In light of this, how can we defend IIA as reasonable for procedures that can pick unique outcomes? We argue that the resolution of this problem is to acknowledge and embrace the fact that no global "best" alternative may exist when making a complex decision (McGann 2006; Moser et al. 2009; Patty and Penn 2008, 2010, 2014). Legitimate procedures can nonetheless pick unique outcomes, and do so in a more meaningful and

[24] We discuss this example in detail in Patty and Penn (2014).

discriminating way than simply breaking ties.[25] The solution lies in expanding the notion of collective choice to include *justification* of the chosen alternative. That is, collective choice through aggregation need not, and arguably should not, be black-boxed in the sense that all that matters is how the alternatives are compared to each other. Rather, particularly when attempting to construct tradeoffs between potentially conflicting criteria, defending one alternative as a final choice may require some explanation or rationale beyond the information embedded in the criteria and alternatives themselves.

This argument is related to, but also diverges from, Sen's argument that choice should depend upon context. We similarly call for accepting what one might term a lower bar of determinacy: faced with a given and fixed set of criteria, it is entirely possible that multiple choices mightbe "uniquely" justified in different ways. However, our argument differs from Sen's because our theory neither relies upon nor contradicts normative desiderata such as fairness, liberty, or rights. Instead, we focus on the dynamic process of constructing a justification for a given decision. As we argue in Patty and Penn (2014), when the principles governing the comparison of alternatives generate an intransitivity, then for any alternative x, some other y dominates it; legitimating the choice of x requires a rationale not only for the choice of x but also for the failure to choose y. We define a notion of legitimacy of a decision sequence characterized by an internal logic and a defensible conclusion.

Finally, Sen's argument (particularly as explicated in Sen 1993) is motivated in part by the idea that the criteria being aggregated might themselves vary with context (e.g., be menu-dependent), whereas our approach takes the criteria as static and fixed. Obviously, Sen's approach is appealing from the standpoint of verisimilitude, but our approach indicates—in line with the spirit of Arrow's theorem itself—that aggregation of multiple criteria is itself sufficient to generate the indeterminacy that Sen seeks to make sense of.

4 Conclusions

All independence axioms are essentially responsiveness axioms. They are desirable in aggregation settings because they dictate that the evaluation of a subset of alternatives be responsive solely to information concerning those alternatives. In this article we have presented a number of practical and theoretical defenses of independence of irrelevant alternatives (IIA) in the context of multi-criterion decision-making, highlighting the fact that a violation of IIA implies that the measurement of the alternatives in set S is not identified without information about alternatives outside of S. At the same time, and as well-known social choice-theoretic results tell us, a fundamental tension exists between the information loss entailed by independence, the desire for responsiveness to multiple criteria in decision-making, and transitivity. This means that the aggregation of multiple and possibly contradictory criteria must be more flexible—and vexing—than the search for a global optimum.

References

Arrow, K. J. (1951). *Social choice and individual values*. New York, NY: Wiley.

[25] Many definitions of legitimacy are possible. A very useful one, which we rely on partially in Patty and Penn (2014), is due to Scott (1995), p. 45, requiring that legitimate procedures reflect "perceived consonance with relevant rules and laws, normative support, or alignment with cultural-cognitive frameworks."

Arrow, K. J. (1959). Rational choice functions and orderings. *Economica*, *26*(102), 121–127.

Arrow, K. J. (1963). *Social choice and individual values* (2nd ed.). New York, NY: Wiley.

Balinski, M. L., & Young, H. P. (2001). *Fair representation: Meeting the ideal of one man, one vote*. Washington: Brookings Institution Press.

Blau, J. H. (1971). Arrow's theorem with weak independence. *Economica*, *38*(152), 413–420.

Blau, J. H. (1975). Liberal values and independence. *The Review of Economic Studies*, *42*(3), 395–401.

Blin, J.-M. (1976). How relevant are "Irrelevant" alternatives? *Theory and Decision*, *7*(1—-2), 95–105.

Bordes, G., & Tideman, N. (1991). Independence of irrelevant alternatives in the theory of voting. *Theory and Decision*, *30*(2), 163–186.

Campbell, D. E., & Kelly, J. S. (2007). Social welfare functions that satisfy pareto, anonymity, and neutrality, but not independence of irrelevant alternatives. *Social Choice and Welfare*, *29*(1), 69–82.

Denicolò, V. (1985). Independent social choice correspondences are dictatorial. *Economics Letters*, *19*(1), 9–12.

Feldman, A. M., & Serrano, R. (2008). ArrowÕs impossibility theorem: Two simple single-profile versions. *Harvard College Mathematics Review*, *2*(2), 46–57.

Fishburn, P. C. (1970). Comments on Hansson's "Group Preferences". *Econometrica*, *38*(6), 933–935.

Fishburn, P. C. (1973). *The theory of social choice*. Princeton, NJ: Princeton University Press.

Gailmard, S., Patty, J. W., & Penn, E. M. (2008). Arrow's theorem on single-peaked domains. In E. Aragones, H. Llavador, & N. Schofield (Eds.), *The political economy of democracy*. Barcelona: UAB.

Gibbard, A. (1973). Manipulation of voting schemes: A general result. *Econometrica*, *41*(4), 587–601.

Gibbard, A. (1974). A pareto-consistent libertarian claim. *Journal of Economic Theory*, *7*(4), 388–410.

Hansson, B. (1969). Group preferences. *Econometrica: Journal of the Econometric Society*, *37*, 50–54.

Hansson, B. (1973). The independence condition in the theory of social choice. *Theory and Decision*, *4*(1), 25–49.

Hylland, A. (1980). Aggregation procedure for cardinal preferences: a comment. *Econometrica: Journal of the Econometric Society*, *48*, 539–542.

Kalai, E., & Schmeidler, D. (1977). Aggregation procedure for cardinal preferences: A formulation and proof of samuelson's impossibility conjecture. *Econometrica*, *45*(6), 1431–1438.

Luce, R. D., & Raiffa, H. (1957). *Games and decisions: Introduction and critical survey*. New York, NY: Wiley.

Mackie, G. (2003). *Democracy defended*. New York, NY: Cambridge University Press.

May, K. O. (1954). Intransitivity, utility, and the aggregation of preference patterns. *Econometrica*, *22*(1), 1–13.

McGann, A. J. (2006). *The logic of democracy: Reconciling equality, deliberation, and minority protection*. Ann Arbor, MI: University of Michigan Press.

Moser, S., Patty, J. W., & Penn, E. M. (2009). The structure of heresthetical power. *Journal of Theoretical Politics*, *21*(2), 139–159.

Murakami, Y. (1968). *Logic and social choice*. London: Routledge & Kegan Ltd.

Patty, J. W., & Penn, E. M. (2018). Definition, measurement, and prescription in politics: Social choice since Arrow. *Annual Review of Political Science* (Forthcoming).

Patty, J. W., & Penn, E. M. (2008). The legislative calendar. *Mathematical and Computer Modelling*, *48*(9–10), 1590–1601.

Patty, J. W., & Penn, E. M. (2010). A social choice theory of legitimacy. *Social Choice and Welfare*, *36*(3), 365–382.

Patty, J. W., & Penn, E. M. (2014). *Social choice and legitimacy: The possibilities of impossibility*. New York, NY: Cambridge University Press.

Patty, J. W., & Penn, E. M. (2015a). Aggregation, evaluation, and social choice theory. *The Good Society*, *24*(1), 49–72.

Patty, J. W., & Penn, E. M. (2015b). Analyzing big data: Social choice & measurement. *PS: Political Science & Politics*, *48*(1), 95–101.

Penn, E. M., Patty, J. W., & Gailmard, S. (2011). Manipulation and single-peakedness: A general result. *American Journal of Political Science*, *55*(2), 436–449.

Ray, P. (1973). Independence of irrelevant alternatives. *Econometrica: Journal of the Econometric Society*, *41*, 987–991.

Riker, W. (1980). Implications from the disequilibrium of majority rule for the study of institutions. *American Political Science Review*, *74*(2), 432–46.

Roberts, K. (2009). Irrelevant alternatives. In K. Basu & R. Kanbur (Eds.), *Arguments for a better world: Essays in honor of amartya sen: Volume I: Ethics, welfare, and measurement*. London: Oxford University Press.

Saari, D. (1994). *Geometry of voting*. Berlin: Springer.

Saari, D. G. (1998). Connecting and resolving Sen's and Arrow's theorems. *Social Choice and Welfare*, *15*(2), 239–261.

Saari, D. (2001). *Decisions and elections*. Cambridge: Cambridge University Press.

Saari, D. (2006). Which is better: The Condorcet or Borda winner? *Social Choice and Welfare*, *26*, 107–129.

Satterthwaite, M. A. (1975). Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, *10*(2), 187–217.

Scott, W. R. (1995). *Institutions and organizations: Ideas and interests*. Los Angeles, CA: Sage Publications.

Sen, A. (2014). Arrow and the impossibility theorem. *The Arrow impossibility theorem* (pp. 29–42). New York, NY: Columbia University Press.

Sen, A. (1970a). *Collective choice and social welfare*. San Francisco: Holden-Day.

Sen, A. (1970b). Interpersonal aggregation and partial comparability. *Econometrica: Journal of the Econometric Society*, *38*, 393–409.

Sen, A. (1970c). The impossibility of a paretian liberal. *Journal of Political Economy*, *78*(1), 152–157.

Sen, A. K. (1971). Choice functions and revealed preference. *The Review of Economic Studies*, *38*(3), 307–317.

Sen, A. (1993). Internal consistency of choice. *Econometrica*, *61*(3), 495–521.

Ubeda, L. (2003). Neutrality in Arrow and other impossibility theorems. *Economic Theory*, *23*(1), 195.

Wilson, R. (1972). Social choice theory without the pareto principle. *Journal of Economic Theory*, *5*(3), 478–486.