Mr. Accuracy, Mr. Simplicity and Mr. Scope: from social choice to theory choice

Michael Morreau
University of Maryland
Center for the Study of Mind in Nature, University of Oslo
mimo@umd.edu

Abstract

Arrow's theorem of social choice (Arrow 1951) has been thought to limit the possibilities for choosing rationally among rival scientific theories on the basis of their accuracy, simplicity, scope and other relevant criteria. It does not. Possible orderings of theories by these criteria are so severely restricted that the theorem is irrelevant in this connection. On the contrary, what is known about social choice in restricted domains implies that there are many acceptable procedures for choosing among theories on the basis of their various merits and demerits.

1. The Analogy between Theory Choice and Social Choice

In scientific research, in engineering and in public policy we sometimes must choose among rival scientific theories. We might hope to do this rationally, first evaluating the alternatives by their accuracy, simplicity, scope and other relevant criteria and then choosing on that basis whichever is, for our purpose, best. One fundamental question is then this: how can we derive an acceptable ordering of rival theories by their <u>overall</u> merit: an ordering that assimilates their various theoretical merits and demerits, and that we can reasonably take to guide our choice between them?

Think of accuracy, simplicity, scope and so on as a society of criteria, each ordering the rival theories by its own lights from better to worse. The question about overall merit then parallels a fundamental question in the theory of democracy: how can we derive a ranking of a society's options that assimilates the various preferences of the individuals that make up the society?

The analogy between social and theory choice runs deeper, in that both proceed under some of the same constraints. In an election, voters do not say how _much_ they like or dislike the candidates. They just rank them first to last, and a social ranking must emerge on that basis. In theory choice too it seems that in important cases we are limited to ordinal information about how rival theories measure up in relevant ways. According to Thomas Kuhn:

> All historically significant theories have agreed with the facts, but only more or less. There is no more precise answer to the question whether or how well an individual theory fits the facts. But questions much like that can be asked when theories are taken collectively, or even in pairs. It makes sense to ask which of two actual and competing theories fits the facts _better_. Kuhn (1970, p. 147)

Another thing that theory choice and social choice have in common is that in both there is a need to balance conflicting claims about what is better and what worse. Once again, a constraint on social-choice procedures applies also to procedures for theory-choice. A _dictator_, in social choice, is someone whose individual ordering of the options is decisive. When he ranks one option higher than another the social ranking always agrees, no matter what anyone else prefers. In theory choice, by analogy, a dictator is a decisive choice criterion. When a theory is better in this one way it is always better overall, no matter how it compares in other ways. Crucially, for what is to come, none of the familiar criteria may be allowed to dictate the comparative overall merit of theories.

You might on empiricist grounds have thought that accuracy is decisive - if not when choosing a theory for some applied purpose then at least in science, when we are after the truth. But accuracy does not dictate overall merit even then. Theories can be brought into agreement with the results of existing observations and experiments by adjusting parameters. Scientific data are generally noisy, though. So if we single-mindedly pursue fit to available data, without regard to other criteria, we will end up choosing overly complicated theories, whose many parameters we can tune to fit the noise. These theories might agree with every error in the data; but they will fit underlying facts and future data less well than do other, simpler theories. They will overfit the data. To avoid this, acceptable choice rules must balance fit against simplicity. And they must do so not in spite of the special importance of accuracy in science but precisely because of it, for balance is what secures accuracy in the long run.[1] Accuracy, then, doesn't dictate overall merit. Plainly, simplicity doesn't either; and nor do scope, fruitfulness or any other familiar criteria.

The theory of social choice has a long history, and has now achieved a high level of technical sophistication in theoretical

[1]I take this point from Forster and Sober (1994).

economics.[2] In a recent article, Samir Okasha raises the intriguing prospect of adapting from this body of knowledge insights into theory choice (Okasha, 2011). His main conclusion, though, is pessimistic. Okasha argues that the theory of social choice establishes a theoretical limit to the possibilities for rationality in science. Specifically, he reinterprets Kenneth Arrow's 'impossibility' theorem (Arrow, 1951) as an argument that there are no acceptable procedures whatsoever for ranking theories by their overall merit, on the basis of ordinal information about how they measure up in the relevant ways.

My first point, here, will be that Arrow's theorem tells us nothing about theory choice on the basis of accuracy, simplicity, scope and so on. Rankings of theories by these criteria are so severely restricted that a crucial assumption of Arrow's is quite unsuitable in this connection. My second point will be that there are in fact many good ways of ranking theories by their overall merit. Other, positive results concerning social choice in restricted domains tell us that quite simple and intuitive procedures are sometimes available, such as counting one theory better than another, overall, if it is better by more criteria than not. The optimistic upshot is that insights from social choice do not so much limit the possibilities for rational choice in science. Rather, they tell us where some of these possibilities lie.

---

[2] Iain McLean and Arnold Urken (1995) have collected contributions to this field from the nineteenth century, the French Enlightenment, the middle ages and ancient Rome. For an overview of more recent work, see Arrow, Sen and Suzumura (2002).

I shall develop the two points as follows. Section 2 sets up Arrow's "multi-profile" framework for studying social choice, and states his impossibility theorem. Section 3 adapts this framework to theory choice, and sets out Okasha's reinterpretation of Arrow's theorem. Section 4 shows that this theorem does not apply to theory choice, by identifying a constraint on procedures for social choice whose analogue in theory choice is absurd. Finally, Section 5 gives several examples of acceptable ways of ranking theories by their overall merit.

2. Arrow's multi-profile framework

Arrow’s "impossibility" theorem answers the question about democratic social choice. The answer is alarming: there is no way whatsoever of assimilating individual into social rankings - none, anyway, that meets what might be taken to be minimal standards of acceptability. This Section sets up Arrow's theoretical framework for studying social choice and states Arrow's theorem.

We will be concerned with choice among some given options on the basis of the preferences of the various members of a society. Individual preferences depend to some extent on personal values, and we are free, within limits, to have a range of these. Accordingly, Arrow's framework allows us to consider social-choice procedures that take into account a more-or-less wide range of possible individual preferences, of preferences that the members of the society can have.

Let $\underline{X}$ be the set of options. A basic assumption is that social choice proceeds on the basis or ordinal information about individual preferences. Accordingly, the preferences of the members $\underline{N}$ =

{1,2,3...n} of the society are represented in this framework by *weak orderings* of X: reflexive, transitive and complete relations. For each $i \in N$, then, there is a range of weak orderings ${}_i{\geq} \subseteq X^2$, the possible orderings of the options by i. A *profile* is an n-tuple $({}_1{\geq}, \ldots {}_n{\geq})$ of individual orderings, one for each $i \in N$. Each profile represents a possible ordering by the entire society, which a social-choice procedure may be required to assimilate into a social ordering. A *domain* D is a set of preference profiles. Notice that in each profile in D, the same individuals N order the same options X. This will be important later on, when we come to adapt this framework in order to study theory choice. A *social-welfare function* f for D is a function mapping each profile in D onto a weak ordering of X.

With Arrow's "multi-profile" framework in place, we turn now to some constraints that might be imposed on social-welfare functions. A first constraint is:

INDEPENDENCE OF IRRELEVANT ALTERNATIVES: For all pairs $S \subseteq X$ and $P_1, P_2 \in D$: if $P_1|S = P_2|S$, then $(fP_1)|S = (fP_2)|S$.[3]

This requires social orderings to supervene on individual orderings. As we move from one profile to another within the domain, there is

[3]For any relation R and set S, the *restriction* of R to S is the relation $R|S$ such that: $x\ R|S\ y$ if $x\ R\ y$ and $x, y \in S$. When $R_1|S = R_2|S$, R_1 and R_2 agree exactly as far as the elements of S are concerned. The restriction $({}_1{\geq}, \ldots {}_n{\geq})|S$ of a profile $({}_1{\geq}, \ldots {}_n{\geq})$ to S is $({}_1{\geq}|S, \ldots {}_n{\geq}|S)$. When $P_1|S = P_2|S$ there is, as we move from P_1 to P_2, no change in the ranking of S by any member of the society.

to be no change in the social ordering of any given pair of options without a change in their ordering by some member of the society.

Another constraint is that whenever everyone ranks one option higher than another, this is decisive. The social ordering always agrees. For any profile $\underline{P}$, let ${}_{i}^{P}\!\geq$ be the $\underline{i}$th component of $\underline{P}$ - this is just $\underline{i}$'s ordering in $\underline{P}$. Writing ${}^{fP}\!\geq$ instead of $\underline{f}(\underline{P})$, this constraint is:

WEAK PARETO: For all $\underline{x},\underline{y}\in\underline{X}$ and $\underline{P}\in\underline{D}$: if for all $\underline{i}\in\underline{N}$, $\underline{x}\ {}_{i}^{P}\!>\ \underline{y}$, then $\underline{x}\ {}^{fP}\!>\ \underline{y}$.[4]

Unanimity might be decisive but, in a democracy anyway, no single individual's ranking is. An individual $\underline{d}\in\underline{N}$ is a <u>dictator</u> of $\underline{f}$ if for all $\underline{x},\underline{y}\in\underline{X}$ and $\underline{P}\in\underline{D}$: if $\underline{x}\ {}_{d}^{P}\!>\ \underline{y}$, then $\underline{x}\ {}^{fP}\!>\ \underline{y}$. A further constraint is:

NONDICTATORSHIP: $\underline{f}$ has no dictator.

A final constraint on $\underline{f}$ concerns the variety to be found among the profiles within its domain. Depending on the options themselves, and on what determines individual orderings, this variety can be quite wide. Where there are no limits to what the individual orderings can be, we may require $\underline{f}$ to have an:

UNRESTRICTED DOMAIN: For all weak orderings ${}_{1}\!\geq,\ldots\ {}_{n}\!\geq\ \subseteq \underline{X}^{2}$: $({}_{1}\!\geq,\ldots\ {}_{n}\!\geq) \in \underline{D}$.

---

[4] Strict orderings $>$ are defined as usual: $\underline{x}>\underline{y}$ if $\underline{x}\geq\underline{y}$ but not $\underline{y}\geq\underline{x}$.

One justification for this constraint is that a social-welfare function, if it is to respect the freedom and sovereignty of the individual members of society, must deliver up a suitable social ordering no matter what their individual orderings turn out to be. Given that these can be anything at all, the social-welfare function had better be ready for everything.

Now Arrow's theorem states that these constraints are incompatible:

ARROW'S THEOREM: Suppose $\underline{n} \geq 2$ and $|\underline{X}| \geq 3$. No social-welfare function satisfies WEAK PARETO, INDEPENDENCE, NONDICTATORSHIP and UNRESTRICTED DOMAIN.

No procedure for aggregating individual orderings of a society's options can meet what might appear to be quite minimal standards of acceptability.[5]

3. Okasha's Arrovian Nihilism

Okasha (2011) argues that analogues of Arrow's constraints are applicable in theory choice on the basis of comparative accuracy, simplicity, scope and other Kuhnian theoretical values (Kuhn, 1977a). The implication is that there is no acceptable way of

[5]For a proof, see a standard text such as (Gaertner, 2009).

comparing rival theories by their overall merit, and therefore no choosing among them so as to maximize overall merit.

3.1 Arrow's Theorem and Theory Choice

Adapting Arrow's framework, we will be concerned with choice among some set X of theories on the basis of a finite set N of choice criteria. We include, among the N, such criteria of accuracy, simplicity and scope as are suitable for evaluating the theories X. For any given set of data, we may assume that there is a corresponding weak ordering ${}_{\text{accuracy}}\geq$ of the theories by their fit to this data. The other criteria, we will assume, induce their own weak orderings ${}_{\text{simplicity}}\geq$, ${}_{\text{scope}}\geq$, etc., of X.

That theoretical orderings are complete might seem far-fetched. We might be able to compare, say, statistical models among themselves with respect to their simplicity (there will be more on this in Section 4). But surely we cannot compare these models in this respect with Darwin's theory of evolution, or with Ptolemaic astronomy. Nor can we compare these two among themselves. That the theoretical rankings are complete is more plausible if the theories are not too different, though. For realism, we may think of X as a smallish set of close rivals.

Continuing now to develop the analogy between Arrovian social choice and theory choice, a *theoretical profile* is a list of possible orderings of X by all the criteria in N: (${}_{\text{accuracy}}\geq$, ${}_{\text{simplicity}}\geq$, ${}_{\text{scope}}\geq$,...). A domain D is a set of profiles. Notice that, true to the analogy, in any profile of D it is the same criteria N that rank the same theories X. Finally, a theory-choice rule for D maps each profile in D onto a weak ordering of X, their comparative overall merit.

Analogues of Arrow's constraints are easily stated. INDEPENDENCE requires the comparative overall merit of any pair of theories to depend entirely on how they measure up by the various relevant criteria. This requirement, Okasha notes, has some intuitive appeal. The analogue of WEAK PARETO requires one theory to be better than another, overall, whenever it is better by every criterion. This seems right. We have already found reason to impose NONDICTATORSHIP, which allows no single criterion to be decisive.

The only constraint of Arrow's theorem that remains is the domain requirement. Okasha takes it on with little comment: ‘[the analogue of UNRESTRICTED DOMAIN] seems unexceptionable - however the theories are ranked by the various criteria, the rule must be able to yield an overall ranking. There should be no apriori restriction on the permissible rankings that are fed into the rule’ (p. 92).

The reinterpretation of Arrow's theorem in theory choice is now very straightforward. It tells us that when there are at least three rival theories and two choice criteria, no rule for choosing among the rivals on the basis of these criteria can meet what might appear to be quite minimal standards of acceptability.

Kuhn famously thought that no "neutral" algorithm dictates rational, unanimous theory choice. Rather, he allowed, there are many algorithms, all of which are compatible with the requirements of rationality, but none of which is uniquely correct. Different scientists, though committed to the same criteria and working with the same evidence, may reach different conclusions about which theory is best (Kuhn, 1977a). Okasha's conclusion likewise implies that there may be no saying, once and for all, which of several rival theories is best; but, as he points out, it goes much further than this by calling into question the very possibility of rational theory choice. If there are many acceptable algorithms, choice can be rational in Kuhn's liberal sense that allows different scientists to use different ones. If on the other hand there are no acceptable algorithms whatsoever then, it might seem, there can no rational way of choosing. Thus 'Kuhn makes rational theory choice look difficult, at least if we cleave to a certain conception of rationality, but Arrow makes it look outright impossible.' (Okasha, p. 94)

This is Okasha's Arrovian nihilism about theory choice. And it is very hard to believe. Surely we sometimes can choose among theories on their merits - even if we cannot always agree among ourselves which is the best one! But Arrow's theorem is rigorous. If we accept that theory-choice rules are bound by analogues of Arrow's constraints, we must also accept the nihilistic conclusion. Somewhere, something will have to give.

3.2 An "Escape" through Cardinal Information?

Okasha considers several "escape routes" from his Arrovian nihilism. The one he finds most promising, following a similar development in the theory of social choice, is to allow not only ordinal but also cardinal information to factor into overall merit. Then not only will it be relevant which of the rival theories are more accurate than which others, and which are simpler, and so on. In addition, information about <u>how well</u> the theories measure up will be taken into account. Okasha gives several examples of theory-choice rules that use this richer cardinal information, from Bayesianism and from model selection in statistics. There is no question that such rules are theoretically important and practically useful.

But much of science is not like this. Kuhn, we have seen, thought that in important cases we are limited to ordinal information about how theories measure up. It is not difficult to see how this might be, even if the criteria in principle allow cardinal measurement. Indeterminacy in the theoretical alternatives themselves might have this effect. Imagine evaluating Darwin’s theory of evolution, sometime around 1859. Even basic concepts, such as that of fitness, remain unclear. Crucial parts, such as an account of the mechanisms of heredity, are missing. Such a theory could be made more complete, by clarifying concepts and filling in details. And it could be completed in many different ways: the more a theory leaves unsettled, the more hypothetical completions it has. Now, even if there is a precise answer to the question how well any one of the many completions fits the facts - or how simple it is, or how great its scope, or what have you - there might be a different answer for each one. In such cases there will be no saying precisely how well the theory itself, with all its imprecision and missing detail, fits the facts. Even so, if it fits better than does its rival on all hypothetical completions of both of them, we can say that, of the two, it fits the facts better.[6]

Revolutionary theories such as Darwin's are not the only ones that are affected. Many other theories also use vague concepts and lack detail, and we can expect to find ourselves limited to ordinal

[6]This is supervaluationism about theoretical merit. Compare Kit Fine's (1975) treatment of linguistic vagueness.

information about them as well. This tends to make Arrovian nihilism all the more serious a threat. Okasha's 'escape route' might be available in routine cases of theory choice, where the theoretical alternatives are so well understood and precisely defined that we can have rich cardinal information about their merits and demerits. But this information will not be available in many other cases.

In fact, as we will now see, Arrow’s theorem does not apply to theory choice on the basis of accuracy, simplicity, scope and the rest: not in the routine cases, not in the revolutionary ones, not if the informational basis for theory choice is rich, and not if it is poor. Arrovian nihilism is a false doctrine. There is no need for any 'escape route' at all.

4. Domain Restrictions in Theory Choice

To see why Arrow's theorem fails to get a grip on theory choice, we must follow the analogy with social choice quite carefully. In Arrow's framework, individual orderings are possible orderings of some given social options by the members of a society. In theory choice, they are possible orderings of some given rival theories by their accuracy, simplicity, scope and so on. What UNRESTRICTED DOMAIN requires is that theory-choice rules shall take as input all "logically possible" lists of orderings of this set of rivals by these choice criteria:

UNRESTRICTED DOMAIN: For all weak orderings ${}_{\underline{accuracy}}\geq, {}_{\underline{simplicity}}\geq, {}_{\underline{scope}}\geq, \ldots \subseteq \underline{X}^2$, $({}_{\underline{accuracy}}\geq, {}_{\underline{simplicity}}\geq, {}_{\underline{scope}}\geq, \ldots) \in \underline{D}$.

To impose this constraint is, as we will see, absurd. A theory-choice rule is a function. It has to produce a suitable ranking of the X for each and every profile in D. But in fact very few weak orderings of X represent genuinely possible theoretical rankings. We might require a social-welfare function to be ready for everything, given that the members of society can order the options any way at all. A theory-choice rule, on the other hand, doesn't have to be ready for everything. Mr. Accuracy, Mr. Simplicity, Mr. Fit and the rest can order the given theories in only a few different ways.

Certainly, there is some variety among theoretical profiles. How well any given theories fit the available data depends on which data are available - on which experiments have been done and which observations made. We may expect a theory-choice rule to come up with a suitable comparison of the rivals by their overall merit no matter what, within reason, the data turn out to be.

But apart from this there is little variety among the profiles admissible to theory-choice rules. The comparative accuracy of theories can turn on the available data, much as the individual ordering of any given member of society can turn on his values. But how theories measure up by other relevant criteria does not depend on anything else. It is just a matter of which theories they are. Orderings by these criteria are the same in all admissible profiles.

Consider comparative simplicity. The Copernican model of the solar system, having done away with the device of equant points, is, in a sense relevant to the choice between them, simpler than the Ptolemaic model. And it is hard to see how it could possibly have failed to be, in this sense, simpler. We could, I suppose, load up a heliocentric model with so many equants that it becomes just as complicated as the Ptolemaic model, or more so. But that would no longer be the Copernican model. Profiles ranking these models by their various merits and demerits had better agree in counting Copernican astronomy simpler, in the relevant sense, than is Ptolemaic astronomy. As far as these theories are concerned, there is but a *single* possible ranking ${}_{\underline{\text{simplicity}}}\geq$, and all admissible profiles must have it in common. It is no good requiring theory-choice rules to reckon with rankings that get things *impossibly* wrong!

For an example from routine science, take model selection. A first step in fitting a curve to available data, say about the relationship between two variables, is to determine the form of the relationship. You do that by selecting a *model*, a class of curves. There is the linear model, LIN, the class of linear curves. There are the parabolic and cubic models, PAR and CUB. Comparative simplicity goes by the number of adjustable parameters. Thus LIN is simpler than PAR, its characteristic polynomial lacking the quadratic term, and PAR is simpler than CUB. Now, these models do not have their adjustable parameters by accident. They are *defined* by their polynomials. Their comparative simplicity therefore could not have been different. All admissible profiles of these models must count LIN as the simplest, followed in turn by PAR and CUB. Once again, there is but a single ordering ${}_{\underline{\text{simplicity}}}\geq$ of the rivals and that is the actual one. There can be no variety among profiles as far as this component is concerned.

It is the same with criteria other than simplicity. We might count the Copernican and Ptolemaic models as having the same scope, in that they address the same topics. Both are models of the solar system. They are about where the heavenly bodies are, and when. Newtonian astronomy, which additionally addresses the physical causes of planetary motion, we will count as having greater scope than they have.[7] And it is hard to see how the three theories can possibly be ranked any other way by their scope, if this is what we mean by 'scope'. The Copernican model surely could have been a bit different in some ways. It could have had more epicycles, or bigger ones. But it couldn't have been anything but a model of the solar system and still have been the Copernican model. Newtonian astronomy couldn't have lacked the theory of gravitation, and still have been the theory that it is. The comparative scope of these theories, like their comparative simplicity, is just a matter of which theories they are. As far as they are concerned, all profiles have the same ranking $_{\underline{scope}}\geq$.

Alternatively, we can understand comparative scope in terms of logical content: "[o]ne might take a theory's 'scope' to be its total set of logical consequences, and the relation 'T_1 has at least as much scope as T_2' to mean that T_2's consequence class is a subset of T_1's." (Okasha p. 91)

Thinking of scope in this way, the variety among possible scope orderings is once again extremely limited. In the case of the statistical models, LIN - or, more precisely, the hypothesis that

[7]I do not mean to suggest that these three theories have been rivals. I consider Newtonian astronomy here to make clear that although this notion of 'scope' does not discriminate between Copernican and Ptolemaic astronomy, the actual rivals, it still is a plausible choice criterion. It does discriminate between some theories.

the best curve is linear, some or other curve within this model - has strictly greater scope than Par. Its logical content is expressed by ∃bc(y=bx+c). This entails ∃abc($y=ax^2+bx+c$), the content of Par, but is not entailed by it. The consequence class of Par is therefore a proper subset of that of Lin. Once again, there is nothing accidental about any of this. Lin and Par couldn't have had other consequence classes than they actually have. Nor could the one class have failed to be included within the other, for inclusion among classes is not a contingent matter. The actual scope ranking of the models is the only possible one. Once again, all profiles have the same ranking $_{scope}\geq$.

We have seen that, on plausible and common conceptions of 'simplicity' and 'scope', there is only one way that theories can be ordered by these criteria. This is so in the case of historically significant theories as well as less exciting ones. Perhaps with other criteria there is more variety among possible orderings. However this may be, the examples we have considered already show that to impose unrestricted domains quite generally on theory-choice rules, as if they cannot otherwise be considered even minimally acceptable, is a mistake. Plainly, no good can come of it. What Arrow's theorem shows is how surprisingly much trouble it can cause.

We might, I suppose, take theoretical profiles to represent not only real possibilities but also some that are merely epistemic - states of affairs with which one might mistakenly reckon, not having noticed that they can never be encountered because they are, in fact, impossible. That might help us to spell out, should we ever wish to do so, what it is for someone to have certain misguided theoretical beliefs: believing, say, that the Ptolemaic model is simpler than the Copernican, or that the parabolic hypothesis is more informative than the linear. But this is no reason to insist on feeding impossibly mistaken beliefs into theory-choice rules. And if for whatever reason someone does insist on feeding them in, and there is trouble, it cannot be the rules that are to blame.

In evaluating possible 'escape routes' from Arrovian nihilism, Okasha discusses domain restrictions due to correlations among rankings by several criteria:

> [A] natural domain restriction would apply if two of the criteria of theory choice exhibit an intrinsic trade-off (or correlation)—for example, if a gain in simplicity always means a loss of accuracy. Then, certain profiles would be impossible, and could be legitimately excluded from the domain of the theory choice rule. (p. 97)

Simplicity and fit exhibit such a trade-off in the case of the statistical models. It can never happen that one of these models is simpler than another, and also fits the data better. The fit of a model to the data is the fit of its best-fitting curve, and the best-fitting linear curve, say, cannot fit any better than the best-fitting parabolic curve because it _is_ a parabolic curve. (These models are nested one inside the other: LIN ⊂ PAR ⊂ CUB, since each can be obtained from the next by setting an adjustable parameter to 0.) A gain in simplicity doesn't always quite mean a loss of accuracy because in the special and, given enough noisy data, extremely rare case where the simpler model fits perfectly, the two must fit equally well. But it does rule out a _gain_ in accuracy.

A further correlation in the example is that between the rankings by simplicity and scope, understood as logical content. These vary together. The simpler of two models is always nested within the more complex and therefore has, of the two, the greater logical content.

As an 'escape route', Okasha finds domain restrictions due to trade-offs unpromising: ‘that such trade-offs always exist does not seem very plausible; and anyway there is no guarantee that the resulting domain restriction would be of the right sort to alleviate the Arrovian impossibility’ (p. 97). He is right to doubt that trade-offs always exist. Simplicity and fit trade off in the above example only because the models are nested; and while theoretical alternatives often are nested in routine science, elsewhere they are not so constrained.[8] There is in general no close logical relationship among rival theories, and one rival can improve on another in several ways at once.

The domain restrictions I have identified are not due to trade-offs though, or to other correlations among criteria. They consist in a lack of variety among possible orderings by individual criteria. Some of the criteria, we have seen, must order the alternatives the same way in each and every profile, so that as far as they are concerned there is no variety at all. Now, not only do these restrictions, unlike trade-offs, occur very generally, in historically significant cases as well as routine ones. In addition, they may be expected to block Arrow's theorem wherever they occur - and along with it, they may also be expected to block improved versions of the theorem that tighten up its domain

[8]For discussion of this point, see Forster (2004).

assumption.[9] Here is why. Let us say that a domain is unrestricted with respect to some given triple of alternatives if for each way of strictly ranking three items, there is some profile within the domain in which this triple is ranked that way. With even a single criterion that orders the alternatives the same way in every profile, there is no triple at all with respect to which the domain is unrestricted. Not only is UNRESTRICTED DOMAIN false in this case. So too are weaker assumptions that replace it in improvements of Arrow's Theorem, such as the assumption that the domain is unrestricted with respect to all triples, and the even weaker assumption that the domain has the 'chain property'.[10] The restrictions I have identified in domains for theory choice are just the thing to make Arrow's theorem and its variants harmless.

We have seen that Arrow’s theorem does not limit the possibilities for theory choice on the basis of their comparative accuracy, simplicity, scope and so on. Theoretical profiles are so severely restricted that the theorem cannot get a grip. It is another thing, of course, to have actual examples of acceptable theory-choice rules. The next Section provides some.

5. There are Acceptable Theory-Choice Rules

Domain restrictions inherent in specific kinds of social alternatives, and in the determination of individual preferences,

[9] Arrow's domain constraint is known to be unnecessarily strong. Some domains, though restricted, still contain enough variety to support a tightened-up impossibility theorem.

[10] See Campbell and Kelly (2002) p. 41 for the chain property, and pp. 50-51 for a version of Arrow's theorem that assumes it.

have been the focus of a great deal of research in recent decades. Quite a lot is now known about how, in spite of general limiting results, they can permit the aggregation of ordinal preferences.[11] This Section makes a start on the project of adapting, from this body of knowledge, insights into the possibilities for choosing among scientific theories on the basis of ordinal information about how they measure up in relevant ways.

I shall now set out several examples of domains for theory choice that are Arrow consistent, in the sense that there are choice rules for them that satisfy analogues of all of Arrow's constraints apart from UNRESTRICTED DOMAIN.

For a first example, let us begin by adding to LIN, PAR and CUB a fourth alternative, by fixing a limited range within which the adjustable parameters of LIN may vary. The resulting subset of LIN, call it SUBLIN, has greater scope than LIN. But it has the same number of adjustable parameters, and so the two are equally simple. SUBLIN makes things more interesting by drawing apart the rankings by simplicity and scope, which otherwise coincide. These rankings are now:

Simplicity: SUBLIN ≈ LIN > PAR > CUB[12]

Scope: SUBLIN > LIN > PAR > CUB

All admissible theoretical profiles of these models have these two orderings in common; any variety among them will have to be in the

[11]For recent surveys of some of what is known about domain restrictions in social choice, see Gaertner (2001, 2002) and Le Breton and Weymark (2011).

[12]$M_1 \approx M_2$ means that both $M_1 \geq M_2$ and $M_1 \leq M_2$.

orderings by fit to available date. The fit of models to any given data never decreases as they become more inclusive. We will consider a domain comprising four profiles, corresponding to the following possible rankings by fit to available data:

Fit: (1) SUBLIN ≈ LIN ≈ PAR ≈ CUB

(2) SUBLIN < LIN ≈ PAR ≈ CUB

(3) SUBLIN < LIN < PAR ≈ CUB

(4) SUBLIN < LIN < PAR < CUB

This domain is Arrow consistent. To verify this, we need a choice rule satisfying analogues of the non-domain constraints: WEAK PARETO, INDEPENDENCE OF IRRELEVANT ALTERNATIVES, and NONDICTATORSHIP. Let us take them in turn.

WEAK PARETO requires that if one model M_1 is strictly better than another, M_2, by all criteria, then M_1 is strictly better than M_2 overall. In light of the tradeoff between simplicity and fit, discussed in Section 2, WEAK PARETO is vacuous in this domain. No model is strictly better than any other by both criteria, in any of the profiles (1)-(4).

INDEPENDENCE OF IRRELEVANT ALTERNATIVES requires that if for each criterion c, the ordering by c of M_1 with respect to M_2 in one profile P_1 is the same as it is in another, P_2, then the comparative overall merit of M_1 and M_2 in P_1 is also the same as it is in P_2.

Since WEAK PARETO is vacuous, it is extremely easy to find a choice rule that satisfies INDEPENDENCE. Even a constant rule – it counts all rivals as equally good, overall, in every profile within the domain – will do. We will have a better example, but this one is not entirely without interest because it also satisfies NONDICTATORSHIP. There is no one criterion such that whenever M_1 is ranked above M_2 by that one criterion, M_1 is always above M_2 in the overall ranking. This is easily seen. Each criterion ranks some model above some other, in some profile. Because the constant rule invariably counts all models equally good, overall, no criterion always 'gets its way'.

Arrow consistency is rather easily established, then. The more interesting question, when WEAK PARETO is vacuous, is whether any non-constant rule establishes it. Here is an example of such a rule:

RULE: For profiles (1) and (2), the overall ordering is the weak ordering $\geq$ such that LIN > PAR > CUB > SUBLIN; for profiles (3) and (4), it is the ordering such that PAR > CUB > LIN > SUBLIN.

For intuitive content, RULE is compatible with a policy of choosing, among the theories that fit the data best, one that is simplest - unless this would mean choosing at either extreme of the ranking by scope.

It is now straightforward to show that this set-up satisfies INDEPENDENCE and NONDICTATORSHIP. Beginning with INDEPENDENCE, notice that since the four profiles differ only in their orderings by fit, what is required is that whenever fit ranks two models the same way in two profiles, the overall ordering is the same as well. By inspection of (1)-(4), it is sufficient that:

(I) The overall ordering of {SUBLIN, LIN} is the same in profile (2) as it is in (3) and in (4); and so is that of {SUBLIN, PAR} and so is that of {SUBLIN, CUB}.

(II) The overall ordeing of {Lin, Par} is the same in (1) as it is in (2), and the same in (3) as it is in (4). Similarly for {Lin, Cub}, and finally:

(III) The overall ordering of {Par, Cub} is the same in profile (1) as it is in (2) and in (3).

Now it is easy to see that Rule satisfies (I) – (III), and thus Independence. That this rule validates NonDictatorship can be seen by inspection of profile (4). There Cub has greater fit than does Par, but of the two Par ranks higher overall. Also, Lin outranks Cub in simplicity and scope, but is outranked by Cub overall. Since as we have seen Weak Pareto holds trivially, the domain comprising profiles (1)-(4) is Arrow consistent.[13]

This example is unrealistic. With enough noisy data, models will practically never fit exactly. No matter which data are available, and no matter which underlying relationship generating it, models with more adjustable parameters will fit better. Ranking (4) by fit will be the only one that is ever observed, and the choice rule will always make the same recommendation: pick Par. But the example does refute Arrovian nihilism about theory choice on the basis of accuracy, simplicity, scope and so on, and that was the point. With this out of the way, let us now turn to a second and more realistic example of an Arrow consistent domain. In this example, the models are ranked by other choice criteria. The new criteria are among those actually used in the field, though, and various theoretical profiles within the domain will commonly be observed. This example

[13]Arrow's theorem as we have seen also requires that the rankings by the various criteria and the overall ranking are weak orderings. This is easily verified.

better illustrates the value for epistemology of concepts and insights from theoretical economics.

There are several widely-used criteria for choosing among models on the basis of some set of available data, among them the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). Both of these criteria reward models for fitting the data, but avoid overfitting by penalizing them for excessive complexity. In practice, I have been told, it is common to use both criteria to score models. When they agree which is best, you choose that one. When they disagree, you might allow another criterion to adjudicate. For example, of the two models they count best you might, depending on your personal bias, choose the one that has the greatest scope. Or you might choose the one that is simplest, or that you think most likely to be correct.

This way of proceeding seems unprincipled. For one thing, the AIC and the BIC are alternatives, embodying the same values of fit to the data and simplicity but differing in their philosophical conceptions. Wouldn't it be better to sort out beforehand which to use? Using both might be a way to avoid answering hard questions before field work, and to speak to supporters of both when reporting results. But then, why adjudicate when the criteria disagree, instead of simply reporting the disagreement? I do not have any justification for all this. Here, I shall just point out that we can at least construe this way of choosing as a case of maximizing rationality. It maximizes the comparative overall merit of models, reckoned by counting one model better than another if it is better by most of the three criteria: the AIC, the BIC and scope.

Rankings by the AIC and the BIC may be expected to have an important formal property. Suppose we fix some set of available data. As we pass through the models from the least to the most inclusive, their fit to this data will in general increase. As a result, both the AIC and the BIC will rank the models higher and higher, but only up to a point. Beyond that, as over-fitting sets in, these criteria will penalize the models for their excessive complexity and rank them lower and lower. The rankings by these criteria, then, may be expected to be single peaked in the sense that there is a single best alternative, and alternatives on the same side of this peak are less good, as they get further away along some linear ordering, on either side of the peak.[14] Here, the criteria are single peaked with respect to the same linear ordering, the inclusion ordering of the models. Notice that the scope ranking is also single peaked with respect to this ordering. The less inclusive a model is the greater its scope, so the peak of the scope ranking is whichever model is least inclusive. Notice as well that for any given data, the AIC and BIC can have different peaks, and that for some data the AIC peak will be less inclusive than is the BIC peak, while for other data the inclusion ordering of the peaks will be the other way around. This will be important later on. The scope ranking of the alternatives, on the other hand, being independent of the available data, is the same in all profiles of the domain.

[14] I assume here, realistically enough, that we are choosing between a finite number of models. This way, we can be sure that each of the criteria counts some or other model as best.

For given data, then, rankings by our three criteria may be expected to make up a single-peaked profile - a profile, each ranking of which is single peaked with respect to a common linear ordering. Let us consider a domain of theoretical profiles, each corresponding to a different set of available data, all of which are single peaked.[15]

Now a well-known result from social choice is relevant to the case. Since the number of criteria is odd, we can aggregate the rankings by simple majority rule. That is, we can count one model at least as good as some rival, overall, if the number of criteria by which it is strictly better than its rival is at least as great as the number of criteria by which its rival is strictly better than it is.[16] The resulting overall ranking will be reflexive, transitive and complete.[17] It is not difficult to see that our scientist in the field, who chooses the common peak of the AIC and the BIC if there is one, but otherwise chooses whichever of the two peaks has the

[15]Choice criteria which like the AIC and the BIC reward models for fitting the data, but penalize them for complexity, might, perhaps, on rare occasions, display a second peak. That might happen if there is some chance bump in noisy data to which an excessively complex model can fit itself so very well as to make up for the penalty it receives for its complexity. For the sake of the example, I assume what I take to be the normal case, which is that the profiles are single-peaked.

[16]Equivalently, we count a model at least as good as a rival, overall, if the number of criteria by which it is at least as good as its rival is at least as great as the number of criteria by which its rival is at least as good as it is.

[17]This result is due to Duncan Black (1948). There is discussion and a proof in Gaertner (2009), pp. 43-47.

greater scope, is picking the maximum by this overall ranking, gotten by putting together the AIC, BIC and scope rankings by majority rule. His way of choosing among models may therefore be seen as a case of maximizing rationality, in which the procedure for ranking the alternatives is quite simple and intuitive. Obviously, WEAK PARETO and INDEPENDENCE are satisfied.

NONDICTATORSHIP is satisfied as well, given a minimum of variety among the theoretical profiles within the domain. It is sufficient that in some profile neither the AIC peak nor the BIC peak has maximal scope, and that these peaks are differently ranked by scope in different profiles: in some profile within the domain, the AIC peak has greater scope than the BIC peak, while in another profile the BIC peak has greater scope than the AIC peak. For example, suppose we are considering just three models M_1, M_2 and M_3, that in some profile we have:

SCOPE: $M_1 > M_2 > M_3$

AIC: $M_2 > M_1 > M_3$

BIC: $M_3 > M_2 > M_1$

while in another profile, with different available data, we have:

SCOPE: $M_1 > M_2 > M_3$

AIC: $M_3 > M_2 > M_1$

BIC: $M_2 > M_1 > M_3$

Then NONDICTATORSHIP is assured. In neither profile do the AIC and BIC peaks have maximal scope. In the first profile, the AIC peak is M_2, which has greater scope than the BIC peak, M_3, while in the second profile the AIC peak is M_3, which is outranked in scope by the BIC peak, M_2. The second profile is the more usual kind with these criteria; since the BIC penalizes models more heavsily for complexity than does the AIC, it tends to select simpler models. But profiles of the first type can be had with very small data sets.

We have seen that our domain is Arrow consistent.

Black's result concerns the case where there is an odd number of voters, but single-peaked domains are also Arrow consistent when it is even. This can be shown by introducing 'phantom' voters, whose ranking of the alternatives is the same in all profiles within the domain.[18] In theory choice the criteria of simplicity and scope are well-suited to this role since, as we have seen, they are bound to rank theories the same way in every profile.

In our example, we might think of the AIC and the BIC as a pair of 'real' criteria augmented by the single phantom criterion of scope. The idea of a phantom has a natural interpretation in this connection. The phantom is a bias in the choice algorithm that becomes visible when the real criteria are at odds with each other. One scientist might be biased towards scope, the other towards simplicity. Other personal biases as well might enter into choice algorithms as phantoms, such as for example a preference for high prior probability, or for conservatism. Thus, in broad support of Kuhn’s (1977a) view of the matter, we see how different scientists can use different ordinal algorithms for theory choice, all of which are minimally acceptable but none of which is uniquely so.

6. Conclusion

Arrow’s ‘impossibility’ theorem has been thought to tell us that there is no acceptable way of ranking scientific theories by their

[18]For discussion, see Le Breton and Weymark (2011), p. 200.

overall theoretical merit, on the basis of their comparative accuracy, simplicity, scope and so on. It tells us no such thing. The possible rankings by these criteria are so severely restricted that Arrow's theorem is irrelevant in this connection. On the contrary, insights from the theory of social choice in restricted domains tell us that, in an important range of cases, there are procedures for putting together theoretical rankings that not only satisfy all of Arrow's non-domain assumptions, but are also quite simple and intuitive.[19]

[19]I thank Carsten Hansen, Thomas Hansen, James Ladyman, Aidan Lyon, Daniel Nolan, Graham Oddie, Samir Okasha and Tom Stoneham for helpful discussion.

References

Arrow, Kenneth 1951: Social Choice and Individual Values. New York: John Wiley.

Arrow, Kenneth J., Amartya K. Sen and Kotaro Suzumura (eds.) 2002: Handbook of Social Choice and Welfare vol. 1. Elsevier: Amsterdam.

Black, Duncan 1948: "On the Rationale of Group Decision Making," The Journal of Political Economy 56: 23-34.

Bossert, Walter and John A. Weymark 2004: 'Utility in Social Choice'. In Salvador Barberà, Peter Hammond, and Christian Seidl (eds) 2004: Handbook of Utility Theory, volume 2. Dordrecht: Kluwer, pp. 1099-177.

Campbell, Donald E. and Jerry S. Kelly 2002: 'Impossibility Theorems in the Arrovian Framework'. In Kenneth J. Arrow, Amartya K. Sen and Kotaro Suzumura 2002, pp. 35-94.

Fine, Kit 1975: 'Vagueness, Truth and Logic'. Synthese 54: 235-259.

Forster, Malcolm R. and Elliott Sober 1994: 'How to Tell When Simpler, More Unified or Less Ad Hoc Theories Will Provide More Accurate Predictions'. British Journal for the Philosophy of Science, 45, pp. 1-35

Forster, Malcolm 2004: 'Chapter 3: Simplicity and Unification in Model Selection'. http://philosophy.wisc.edu/forster/520/Chapter%203.pdf

Gaertner, Wulf 2001: Domain Conditions in Social Choice Theory. Cambridge: University Press.

-2002: "Domain Restrictions," in K.J. Arrow, A.K. Sen and K. Suzumura, eds., Handbook of Social Choice and Welfare, Vol. 1, North-Holland: Amsterdam, pp. 131-170.

-2009: A Primer in Social Choice Theory, revised edition. Oxford: University Press.

Kuhn, Thomas 1970: The Structure of Scientific Revolutions, second edition. Chicago: University of Chicago Press.

-1977a: ‘Objectivity, Value Judgment, and Theory Choice’, in his 1977b, pp. 320-339

-1977b: The Essential Tension. Chicago: University of Chicago Press.

Le Breton, Michel and John A. Weymark 2011: “Arrovian Social Choice on Economic Domains,” in K.J. Arrow, A.K. Sen and K. Suzumura, eds., Handbook of Social Choice and Welfare, Vol. 2, North-Holland: Amsterdam, pp. 191-299.

McLean, Iain and Arnold B. Urken (eds.) 1995: Classics of Social Choice. Ann Arbor: University of Michigan Press.

Okasha, Samir 2011: ‘Theory Choice and Social Choice: Kuhn versus Arrow’. Mind 120, 477, pp. 83-115.