Aggregating Out of Indeterminacy: Social Choice Theory to the Rescue

By: Brian Kogelmann

"When sincere and good persons differ, we are prone to think they must accept some procedure to decide their differences, some procedure they both agree to be reliable and fair. Here we see the possibility that this disagreement may extend all the way up the ladder of procedures."

- Robert Nozick, *Anarchy, State, and Utopia*, pp. 98

§1 Introduction

Public reason liberalism says that moral rules must in some distinctive sense be justified to all. Gerald Gaus's *The Order of Public Reason* is the most sophisticated account of public reason liberalism that has been offered since John Rawls's *Political Liberalism*. Gaus believes there is indeterminacy as to which moral rules are justified to all: rules r_1, r_2, and r_3 might all be justified to Althea, Bertha, and Cassidy respectively. Yet in order to have social order we need one and only one rule to govern social life: either r_1, r_2, or r_3. Since all three rules are justified to all, how do we select which rule should ultimately be implemented? The problem is more pronounced when we allow for the possibility of Althea, Bertha, and Cassidy to have diverging preferences over r_1, r_2, and r_3. Given that Althea most prefers r_1, Bertha most prefers r_2, and Cassidy most prefers r_3, which rule should we adopt? Call this the *indeterminacy problem.*

One solution to the indeterminacy problem offered by some public reason liberals is to use a social choice mechanism to choose between r_1, r_2, and r_3. Gaus does not think this is a viable solution to the indeterminacy problem (Gaus 2011: 327-330). Instead, Gaus offers a unique way of solving the indeterminacy problem. He argues that an evolutionary equilibrium selection mechanism – the details of which we discuss in the next section – will ultimately lead Althea, Bertha, and Cassidy to settle on one and only one rule (Gaus 2011: chp. 7). After highlighting a criticism of Gaus's solution to the indeterminacy problem in the existing literature, I argue that, *contra* Gaus's initial concerns, with the tools of social choice theory we *can* select a uniquely justified rule from r_1-r_3. Moreover, we can do so in a way that avoids the worries raised against Gaus's solution to the indeterminacy problem as well as the worries Gaus himself raises against the use of social choice mechanisms. Furthermore, we can accomplish this in a manner that more reliably yields a publicly justified outcome when compared to Gaus's own model of public reason. Social choice theory thus rescues public reason liberalism by aggregating out of indeterminacy.

§2 Public Reason, the Indeterminacy Problem, and the Failures of Social Choice

At the core of the public reason project is a search for rules that are justified to all. In commencing this search some level of idealization is applied to those persons that make up society. One feature of Gaus's version of public reason liberalism is the minimal level of idealization he places on persons – there is still a good deal of diversity concerning the evaluative standards of the Members of the Public, what Gaus calls the class of idealized individuals we seek to justify moral rules to. Because there is a diversity of evaluative standards, Members of the Public will disagree on which moral rules are justified and, even when they do agree on which moral rules are justified, they will disagree on how to rank them: "Given deep pluralism of the evaluative standards on the basis of which each ranks the proposals, there will be similar pluralism in the rankings of the proposals" (Gaus 2011: 310). This presents a problem: when Members of the Public disagree on which moral rules are justified and how we ought to rank these rules, how do we choose which rule should actually be implemented to govern social life? This is the indeterminacy problem, for it is indeterminate which rule we should ultimately select. This would not be so if Members could agree on one and only one rule as best.

To make things more concrete let us consider an example. Suppose we have three Members of the Public: Althea, Bertha, and Cassidy. Suppose there are five rules up for debate, r_1-r_5. Suppose our three parties rank the rules as shown in Table 1. In all of the rankings is the option of "blameless liberty," in which no moral rule governs the behavior in question. If some Member of the Public ranks a rule below blameless liberty then that rule is not justified. The reason why Gaus uses blameless liberty as the relevant standard has to do with the goals of his project: he inquires whether we are justified when we make claims on others; such claims are justified on his view just in case persons would endorse such restrictions on their actions. If a person places a rule *above* blameless liberty then it means they have some kind of reason to endorse that rule, since they would rather have that rule when compared to no rule at all; if a person places a rule *below* blameless liberty then it means they do not have any reason to endorse that rule, for they would rather have no rule governing that particular area of social life when compared to the rule in question (Gaus 2011: 319).

Althea	Bertha	Cassidy
r_1	r_1	r_2
r_2	r_2	r_1
r_3	r_3	r_3
r_4	r_4	r_5
Blameless Liberty	Blameless Liberty	Blameless Liberty
r_5	r_5	r_4

Table 1

Taking the set of all rules placed above blameless liberty by every Member of the Public gives us an *eligible set* of justified moral rules – eligible set, for all the rules in the set are eligible for implementation in the sense that each rule in the set is justified to all. In Table 1, the eligible set consists of rules r_1-r_3. To be justified, the moral rule that ultimately ends up governing social life must come from this set. But which moral rule should we choose? Here Gaus further narrows down candidates for public justification by applying the Pareto rule to the eligible set. This gives us an *optimal eligible set*, consisting of r_1 and r_2. But even after reaching the optimal eligible set we still lack one single rule to govern social life: should we have rule r_1 or r_2? We thus face the indeterminacy problem.

One way of narrowing down the optimal eligible set to a singleton is by using a social choice mechanism. A social choice mechanism is a way of generating either an ordering of options or a single best option by taking individual preference orderings over available options as inputs, and having either an ordering or a single choice as the output, where this output is determined by a set of aggregation rules applied to the individual preference orderings taken as the input. Systems of voting are quintessential examples of social choice mechanisms. To continue the example, a social choice mechanism would take Althea, Bertha, and Cassidy's preferences over r_1 and r_2 (because this is the optimal eligible set) and, through the application of aggregation rules on these orderings, select one moral rule to govern social life.

Gaus presents two problems for the use of social choice mechanisms to solve the indeterminacy problem. The first problem is that there is no unequivocally best social choice mechanism to use (Gaus 2011: 327-328). Relevant here is Kenneth Arrow's impossibility theorem, which shows that there is no social choice mechanism satisfying four intuitively plausible axioms that any aggregation rule ought to satisfy (Arrow 1951/2012). Because Arrow's axioms are plausibly the best hope of composing the uniquely best social choice mechanism, and because no social choice mechanism can simultaneously satisfy these axioms, there is no uniquely best mechanism we can apply to Althea, Bertha, and Cassidy's preferences.

Absent a uniquely best social choice mechanism, one way of still using social choice theory to solve the indeterminacy problem is by appealing to the social choice mechanism Althea, Bertha, and Cassidy all think is best. This seems the appropriate way of proceeding regardless. After all, *even if* Arrow's impossibility theorem was false, and *even if* Arrow's axioms implied one and only one social choice mechanism, it would seem at odds with public reason liberalism's stated project to impose such a social choice mechanism on Althea, Bertha, and Cassidy if all three believed there was something deeply misguided with aggregating preferences in such a manner – indeed, this would be akin to imposing a moral rule on Althea, Bertha, and Cassidy all place below blameless liberty. A natural way around this worry is to ask which social choice mechanism Althea, Bertha, and Cassidy think is best, and, from there, using that unique aggregation rule to transform the three preference orderings over the optimal eligible set into one unique moral rule to be realized.

Here is where Gaus's second criticism comes in (Gaus 2011: 329). Just as our lack of idealizing Members of the Public led to conflicting preferences over moral rules and thus indeterminacy as to what is justified, it is likely that evaluative diversity among Althea, Bertha, and Cassidy will lead to conflicting preference orderings over which social choice mechanism should be employed.[1] We thus have a meta-indeterminacy problem, displayed in Table 2. To illustrate, consider

[1] It is of course possible that a Member of the Public will think that *no* social choice mechanism is an appropriate way of resolving the indeterminacy problem; this would be akin to a Member of the Public thinking *no* moral rule for some particular circumstance is better than blameless liberty. In these kinds of cases, Gaus's version of public reason liberalism falls silent, in that it admits that no

five social choice mechanisms: plurality voting, plurality runoff, the Condorcet method, approval voting, and the Borda count. As was the case with rules r_1-r_5, our Members of the Public have diverging preferences over which social choice mechanism is best. In such a case, what social choice mechanism should we use to solve the original indeterminacy problem over rules r_1 and r_2? As Gaus notes: "evaluative diversity will again manifest itself, and Members of the Public will arrive at different orderings of the possible selection procedures" (Gaus 2011: 329). Call this meta-indeterminacy problem *Nozick's objection*, based on the epigraph of this paper.

Althea	Bertha	Cassidy
Plurality Voting	Condorcet	Borda Count
Plurality Runoff	Borda Count	Condorcet
Condorcet	Plurality Voting	Approval Voting
Approval Voting	Approval Voting	Plurality Voting
Borda Count	Plurality Runoff	Plurality Runoff

Table 2

In response to Nozick's objection, how does Gaus solve the indeterminacy problem and narrow down the optimal eligible set to a singleton? Instead of finding a rational procedure (like the use of a social choice mechanism or some kind of game theoretic bargaining model) to justify the rule that is ultimately selected, Gaus relies on an evolutionary selection mechanism. Actual persons in society, through making choices, can end up coordinating on one moral rule through bandwagon effects, increasing returns, and path-dependency. This one moral rule that citizens coordinate on through such an evolutionary process is ultimately justified (so long as it is in the optimal eligible

---

rule in this particular case and circumstance is publicly justified. Similarly, the social choice solution to the indeterminacy problem falls silent if some think that no aggregation procedure should be employed. I thank an editor at *Politics, Philosophy, & Economics* for raising this possibility.

set), because "*consulting simply his or her own evaluative standards, each has decisive reason to freely endorse whichever moral requirement they have coordinated on*" (Gaus 2011: 394). Continues Gaus: "Once society has chosen a rule, if the rule in equilibrium is also a member of the optimal eligible set, *we have created through our actual independent choices what impartial reason could not deliver: a unique justified rule*." (Gaus 2011: 402).

Fred D'Agostino (2013: 145-153) criticizes Gaus's reliance on an evolutionary selection mechanism to solve the indeterminacy problem by focusing not on the *outcome* of the evolutionary process, but on the actual evolutionary *process* itself. According to D'Agostino, Gaus places the following two conditions on his solution to the indeterminacy problem. If both conditions are satisfied, then Gaus considers his solution to be a success. These conditions are:

> (a) The mechanism takes us to a single determinate set of rules.

and

> (b) Each citizen, consulting her own evaluative standards, has reason to follow the set of rules.

These two conditions only appraise the outcome of the mechanism. But if one is *also* concerned with the process of the mechanism itself, then here is another plausible condition one might place on the mechanism:

> (c) The mechanism takes us to a single determinate set of rules in a manner that secures the moral status of the resulting equilibrium.

The question is whether Gaus's solution to the indeterminacy problem satisfies (c). Here, D'Agostino does not claim that the evolutionary solution necessarily violates (c). Rather, he claims that many instances of social evolution rely on bargaining asymmetries, bribery, intimidation and the like, all of which might cast doubt on whether the resulting equilibrium comes about in a manner that strikes us as appropriate from the moral point of view.[2] When this happens Gaus's evolutionary solution does not satisfy (c).

I do not want to debate D'Agostino's criticism of Gaus. It seems reasonable to me, so we shall grant that (c) is a plausible condition solutions to the indeterminacy problem should satisfy, and that Gaus's solution sometimes runs afoul of (c). But then we must accept that, as solutions to the indeterminacy problem, *both* social choice mechanisms *and* Gaus's evolutionary selection mechanism fail. Many times Members of the Public will have conflicting preferences over social choice mechanisms, leading to Nozick's objection. Many times social evolution will involve morally suspect factors. In light of this, D'Agostino is led to embrace what we shall call *Rawls's conclusion*: "It may turn out that, for us, there exists no reasonable and workable conception of justice at all" (Rawls 1980/1999: 356). Let us see if social choice theory can rescue us from Rawls's conclusion.

§3 Rehabilitating Social Choice

3.1 *Not the Same Social Choice Rule for All, Just Some Social Choice Rule for All.*

In responding to Nozick's objection the first thing to note is that there is no reason *one and only one* social choice mechanism must be applied to the optimal eligible set. When confronted with the optimal eligible set the problem we face is how to narrow down to a singleton in a manner that treats all as free and equal moral persons. If multiple social choice mechanisms can be applied to the

[2] Indeed, Bowles and Gintis (2011: chp. 8) show that the evolution of altruistic sentiments in human beings was largely grounded in warfare, particularly the ability of altruistic groups to exterminate competing groups that lacked such sentiments.

optimal eligible set and still yield a singleton then we have solved the problem.[3] Of course, the worry is that different social choice mechanisms applied to the same optimal eligible set will yield different singletons, a problem we shall address down the road. But for now, we note that it *could* be the case that application of Althea's most preferred social choice mechanism and application of Bertha's most preferred social choice mechanism select the *same* moral rule. In such a case there seems to be nothing wrong with applying both mechanisms. We apply Althea's most preferred social choice mechanism to the optimal eligible set and, in doing so, yield singleton r_2. By virtue of applying Althea's most preferred aggregation rule we treat her as a free and equal moral person. We then apply Bertha's most preferred social choice mechanism to the optimal eligible set and, in doing so, also yield singleton r_2. Because we have applied Bertha's most preferred aggregation rule we also treat her as a free and equal moral person. Call this the *convergence solution* to Nozick's objection.

Besides the worry that social choice mechanisms will not all converge in such a serendipitous manner, a more foundational objection to the current proposal is that Members of the Public might object to applying certain social choice mechanisms preferred by other Members to the optimal eligible set. Suppose Althea's most preferred social choice mechanism is plurality runoff, that Althea thinks Bertha's most preferred social choice mechanism is a deeply objectionable way of aggregating

---

[3] Of course, this assumes that selecting a rule Althea endorses (places above blameless liberty) with the social choice mechanism she thinks is best yields a publicly justified outcome in a way that strikes us as appropriate from the moral point of view (something Gaus seems to agree with, as we saw in the section above when discussing his second criticism of the social choice solution). If one denies this then applying multiple social choice mechanisms to solve the indeterminacy problem will not seem a compelling solution. I thank an editor at *Politics, Philosophy, & Economics* for pointing out this underlying assumption of my argument.

preferences, and that both methods indeed select the same moral rule. Do we fail to treat Althea as free and equal when we *also* apply Bertha's most preferred social choice mechanism to the optimal eligible set? If so then we might think that something has gone wrong in how we narrowed down to a singleton. Though we narrow down to one moral rule, we might have violated criterion (c) for, even though we applied Althea's most preferred social choice mechanism to the optimal eligible set, Althea thinks Bertha's social choice mechanism is deeply objectionable and, by the convergence solution, we have also applied Bertha's most preferred social choice mechanism as well.

In addressing this worry we must ask: why would Althea object to Bertha's most preferred social choice mechanism in such an extreme way? Now certainly Bertha's most preferred social choice mechanism might not be Althea's most preferred, just like Bertha's most preferred moral rule might not be Althea's most preferred. But that is not what is at issue here. The worry is that Althea thinks Bertha's social choice mechanism is *so* objectionable (due to some kind of non-attitudinal fact(s) about this rule) that application of that mechanism to the optimal eligible set somehow fails to secure the moral status of the single moral rule eventually chosen, should the mechanisms indeed all converge on the same moral rule. Just as Althea places moral rules below blameless liberty only when they are deeply objectionable (there are many rules she can live with that are not her most preferred), the worry is that Althea thinks similarly of Bertha's favored aggregation procedure (though there are many social choice mechanisms she can live with that are not her most preferred, Bertha's is not one).

I can think of three reasons why Althea might have such strong objections: (*i*) Bertha's most preferred social choice mechanism might unjustly favor Bertha over all other Members of the Public; (*ii*) Bertha's most preferred social choice mechanism could be such that it fails to actually select a moral rule, either because it creates a cycle or a tie; or (*iii*) Bertha's most preferred social choice mechanism might be manipulable. Barring (*i*)-(*iii*), sincere persons who desire to engage in the practice of justifying a shared social morality would likely not in good faith object to the use of a social choice mechanism as so misguided that it somehow taints whatever rule it ultimately selects, thereby violating criterion (c). I now argue that concerns (*i*)-(*iii*) are ruled out by the formal features of Gaus's justificatory apparatus. That is, given the assumptions of Gaus's overall framework which we work within, we do not need to worry about (*i*)-(*iii*). The objection raised against the convergence solution in this section thus fails.

With worry (*i*) Althea is concerned that Bertha's most preferred social choice mechanism will unjustly favor Bertha. Here is one such mechanism Bertha could choose as her most preferred: Bertha's preference ordering over the optimal eligible set determines the collective ordering over the optimal eligible set for all Members of the Public. Thus, if Bertha most prefers r_1 then the application of her social choice mechanism to the optimal eligible set will yield singleton r_1. Now suppose Althea's most preferred social choice mechanism also chooses rule r_1 from the optimal eligible set. The convergence solution to Nozick's objection thus apparently succeeds in solving the indeterminacy problem. Yet we might think that we have solved the indeterminacy problem in a morally unsatisfactory way in the specific case under discussion by virtue of employing Bertha's utterly selfish aggregation rule. More practically, we might worry that if Members of the Public are allowed to employ social choice mechanisms like the one employed by Bertha then the convergence solution is unlikely to ever succeed: for every Members' most preferred social choice mechanism will just select their most preferred moral rule. Then, so long as the indeterminacy problem is actually a problem, the convergence solution trivially fails in reducing to a common singleton.

In addressing Althea's worry we appeal to the formal constraints Gaus places on the moral rules proposed by Members of the Public: proposed moral rules must satisfy these formal constraints before they can even be permissible candidates for public justification in the first place. Following Gaus's lead, we insist that these same formal constraints *also* be placed on the social choice mechanisms employed by Members of the Public. As Gaus notes, such "constraints are not ad hoc: they specify features that proposals must have for you and me to consider them as serious candidates for a justified rule in social morality" (Gaus 2011: 294). Again following Gaus, the constraints we place on social choice mechanisms are not ad hoc, for they specify features social choice mechanisms must satisfy in order for you and me to consider the application of these mechanisms to be morally justified solutions to the indeterminacy problem, and thus morally justified solutions to the problem of a justified shared social morality.

In a Rawlsian spirit one constraint Gaus places on moral rules Members of the Public may propose to one another is that such rules must satisfy *generality*: "That is, it must be possible to formulate them without the use of what would be intuitively recognized as proper names or rigid definite descriptions" (Rawls 1971: 131). In a similar vein we insist that social choice mechanisms also satisfy generality to ensure that our solution to the indeterminacy problem ends up solving the problem in a way that strikes us as appropriate from the moral point of view. Formally, we stipulate that all social choice mechanisms that Members of the Public may employ must satisfy *anonymity*, which requires the output of a social choice mechanism be invariant across permutations of names and orderings. Clearly, Bertha's proposed social choice mechanism violates anonymity. Indeed, anonymity effectively placates Althea's first worry. Since social choice mechanisms must be invariant across permutations of names and orderings, it is impossible for Members to choose a social choice mechanism that unduly favors themselves or certain groups. This takes care of Althea's first reason for objecting to application of Bertha's most preferred social choice mechanism to the optimal eligible set: such aggregation rules are simply off-limits.

Now consider worry (*ii*). Here, Althea objects to the application of Bertha's most preferred social choice mechanism because such a mechanism may fail to select a moral rule at all, either because the mechanism creates a cycle or tie. Intuitively we know what a tie is: a social choice mechanism produces a tie in our context if it says both r_1 and r_2 are the best elements of the optimal eligible set. And as we know from the Condorcet paradox, aggregation procedures can induce cycles: they can say that r_1 is better than r_2 which is better than r_3… which is better than r_1 again, and so on and so forth. A legitimate worry Althea might have is that Bertha's most preferred social choice mechanism might do either of these two things – if it does, then it fails to solve the indeterminacy problem at all, for it fails to reduce the optimal eligible set to one single moral rule.

As before, extending Gaus's formal constraints on proposed moral rules can mollify Althea's worry. One such constraint is that proposed moral rules must actually provide some kind of resolution to the problems they are introduced to be solutions to: if Althea wants to φ in circumstance C, yet Bertha wants to Ψ in circumstance C, where φ-ing and Ψ-ing are incompatible, then a moral rule governing C must actually tell us whether Althea is permitted to φ or Bertha is permitted to Ψ (Gaus 2011: 298). That is, moral rules must actually perform the function we introduce them to perform. Since the point of introducing social choice mechanisms in the first place is to reduce the optimal eligible set to a singleton, we similarly stipulate that all such mechanisms must actually do this: all social choice mechanisms applied to the optimal eligible set must reduce the optimal eligible set to one and only one moral rule. We do this by stipulating that social choice mechanisms capable of inducing a cycle or a tie come with some kind of cycle- or tie-breaking rule.

The final worry (*iii*) is that Bertha will choose a social choice mechanism that is manipulable. By this we mean that the social choice mechanism is such that some Members of the Public will have incentive to misrepresent their preferences in order to get a moral rule selected they more prefer than the moral rule that otherwise would be selected were they to elicit their honest preferences. And, as the Gibbard-Satterthwaite theorem shows, all non-dictatorial social choice mechanisms can be manipulated in such a way (Gibbard 1973; Satterthwaite 1975). Since dictatorial aggregation procedures are ruled out by our requirement of anonymity, it follows that all permissible mechanisms Members may use will be manipulable. In response to this genuine worry we stipulate that Members of the Public do not misrepresent their preferences over the optimal eligible set – they always state what their honest preferences are.[4]

Though this stipulation does not correspond to a formal constraint Gaus places on the moral rules proposed by Members of the Public, it does relate to a more general feature of his justificatory apparatus: "In judging whether to endorse a proposed rule, our Members of the Public rely only on the reasons they *have* to endorse it. They do not bargain, bluff, or engage in strategic behavior: you and I are not interested in what bargains Members of the Public might strike, but the reasons they recognize, because those are the reasons that actual agents have" (Gaus 2011: 276). In continuing to model Members of the Public as genuinely reflecting on the reasons they have to endorse proposed moral rules we idealize away the possibility of manipulating social choice mechanisms. As such, though Bertha's most preferred social choice mechanism will certainly be

[4] Adding this formal constraint grants that Althea is correct in her worry about the objectionable properties of manipulation. Another response here is to argue that Althea is incorrect about this: that manipulation is a *good* thing, in that it has desirable properties. See Dowding and van Hees (2008) for argument along these lines. If one takes this route then, obviously, one would not add the corresponding formal constraint we have just introduced. I thank an anonymous reviewer at *Politics, Philosophy, & Economics* for suggesting this possibility.

manipulable, this is not something Althea need worry about. We thus see that all considerations (*i*)-(*iii*) Althea might have for objecting to the use of Bertha's most preferred social choice mechanism fall by the wayside; simply extending the formal features of Gaus's justificatory framework deals with these worries.

3.2 *The Discursive Dilemma.*

Now it might be argued that the convergence solution can still run afoul of criterion (c) – in that the convergence solution solves the indeterminacy problem in a morally suspect way – for reasons having to do with the *discursive dilemma*, the central impossibility result from the theory of judgment aggregation (List and Pettit 2002).[5] To illustrate the worry, suppose that Althea's most preferred social choice mechanism is Condorcet, Bertha's is Borda, and Cassidy's is plurality runoff; further suppose that all three aggregation procedures select from the optimal eligible set rule r_2 to be realized. Now consider Table 3. Here, our three Members of the Public are posed two questions: first, whether the particular social choice mechanisms employed give them reason to support the rule selected; and second, whether they have reason to support rule r_2 as the selected rule.

[5] I thank an anonymous reviewer at *Politics, Philosophy, & Economics* for raising this objection.

	Do the following social choice rules give you reason to support the rule r_2 they ultimately select?			Do you have reason to support the selected rule, r_2?
	Condorcet	Borda	Plurality Runoff	
Althea:	Yes	No	No	Yes
Bertha:	No	Yes	No	Yes
Cassidy:	No	No	Yes	Yes
Majority:	No	No	No	Yes

Table 3

Now because each Member's most preferred social choice mechanism by hypothesis selects r_2 all hold they have reason to support r_2 as the selected rule. But as one can see, a simple majority all believe that employment of each social choice mechanism does not give one reason to support the rule ultimately selected, *even though* all hold they have reason to support the rule that *is* ultimately selected from the application of all such mechanisms. In this case the convergence solution has worked – we apply each Member's most preferred aggregation procedure and, as a result, converge on the same rule – yet intuitively we have still run afoul of criterion (c), for there seems to be something suspect about applying a series of social choice mechanisms that are each disapproved of by a majority of Members of the Public, even though that is what the convergence solution calls for in this particular case.

In addressing this objection we need more detail concerning what it means for a Member of the Public to answer "yes" or "no" to the question: "does the following social choice rule give you reason to support the rule r_2 it selects?" On one understanding of what this could mean, Members answer "no" to this question just in case the mechanism is not their most preferred mechanism, and "yes" to this question just in case the mechanism is their most preferred mechanism. But if this is how we interpret it, then Table 3 poses no real challenge at all. For on this understanding, Table 3 merely tells us that there is no social choice mechanism most preferred by a majority of Members; yet still, all Members' most preferred mechanism selects r_2. Certainly a state of affairs of this kind is not troubling enough to think the convergence solution, as applied to situations like Table 3, runs afoul of (c), for there is nothing morally questionable about citizens disagreeing about what the most reasonable way of aggregating preferences is, which is what Table 3 says on this interpretation of how the citizens answer the question under consideration.

Perhaps a more charitable interpretation of what it means to answer "yes" or "no" to our question – "does the following social choice rule give you reason to support the rule r_2 it selects?" – is that Members answer "no" when they think the social choice mechanism is so objectionable a method of aggregation that it taints whatever rule it ultimately selects, and that Members answer "yes" when they think otherwise – that even though the mechanism might not be their most preferred mechanism, it is still a reasonably fair way of aggregating preferences and solving the indeterminacy problem. If this is how we understand the situation posed in Table 3 then we again have nothing to worry about. Though this state of affairs *would* actually be worrying (in that here it *does* seem intuitive to say that (c) has actually been violated), such a state of affairs cannot actually arise, for reasons already given in §3.1.

In §3.1 it was argued that, when we restrict our attention to persons with a sincere desire to engage in the practice of justifying a shared social morality, plausible reasons such a person could have for deeming a social choice mechanism to be *so* objectionable that it irredeemably taints the resulting selected rule is preempted by simply extending the formal features of Gaus's justificatory apparatus: (*i*) dictatorial rules are off-limits given the requirement that all rules satisfy generality and thus anonymity; (*ii*) all rules come with a cycle- or tie-breaking mechanism; and finally (*iii*) we stipulate that sincere preferences are always given. Since these are the three main concerns a reasonable person could raise against a preference aggregation mechanism as being so misguided that it somehow taints whatever rule it ultimately selects, it follows that Members of the Public cannot object to the employment of other Members' most preferred social choice mechanisms in such a way. But if this is true, then, given the current interpretation of what it means for a Member to think the application of a particular social choice mechanism fails to be reason-giving, situations like Table 3 simply cannot arise.

Now it could be that citizens have other reasons besides (*i*)-(*iii*) for objecting to social choice rules as being so objectionable that they fail to be reason-giving in any sense of the term. Whether the formal features of Gaus's justificatory apparatus can be extended to deal with these issues is a question that must be answered on a case-by-case basis – it cannot be addressed *a priori*. Still, the analysis in §3.1 showed that three common objections good-willed persons committed to justifying a shared social morality could raise against the employment of certain social choice mechanisms are ruled out. We can thus tentatively state that, if this is how we interpret what it means for a social choice mechanism to give someone a reason to support the rule it selects, then situations like Table 3 cannot happen. To conclude, on two plausible interpretations of what is actually happening in Table 3, either situations like Table 3 do not run afoul of (c), or situations like Table 3 cannot occur. Either way, problems arising from the discursive dilemma do not threaten the convergence solution.

3.3 *Riker's Challenge.*

Having shown that the convergence solution does not run afoul of (c) in the prior two sections, we adequately address Nozick's objection and silence Rawls's conclusion if the differing social choice mechanisms of the varying Members of the Public applied to the optimal eligible set select the same moral rule: in such a case, we solve the indeterminacy problem in a manner that treats all as free and equal moral persons. But what is the likelihood of this happening? If it is likely to happen then we can be optimistic about solving the indeterminacy problem in a manner satisfying (c), resulting in publicly justified outcomes that strike us as appropriate from the moral point of view; if not, then, given the failures of Gaus's evolutionary solution, it is unclear where the public reason liberal turns to show how a justified social morality can come about in a manner that secures the moral status of the resulting equilibrium. The question of the likelihood of the convergence solution succeeding brings us to William Riker's *Liberalism Against Populism*. Riker argues that social choice theory spells trouble for populist justifications of democracy – justifications of democracy grounded in the idea that voting procedures reveal the will of the people. One of the arguments for this conclusion relates to the challenge just posed: Riker argues that social choice mechanisms generally speaking cannot reveal the will of the people, because "different methods of election and committee decision-making have produced different results from identical distributions of preferences" (Riker 1982: 22).

Riker's challenge is premised on the obviously true claim that it is logically possible for different social choice mechanisms applied to the same optimal eligible set to choose different moral rules as best. But is this enough to defeat the convergence solution to Nozick's objection? It can't be. If logical possibility was sufficient to make us throw out the convergence solution then we would be in even bigger trouble than lacking a solution to the indeterminacy problem: for on Gaus's model of how we trim down to the eligible set in the first place, it is logically possible that the eligible set is the empty set. This occurs if each candidate moral rule is placed below blameless liberty by some Member of the Public. Yet such logical possibilities are not enough for Gaus to throw out the entire public reason liberalism project. Though Gaus recognizes that the eligible set may indeed be empty, he gives reason for why this is unlikely: "Denying moral authority to basic moral rules is a grave matter indeed. Members of the Public would do so only when confronted by rules that, in their view, palpably fail to adequately perform their tasks" (Gaus 2011: 323). Because these sorts of possibilities are unlikely due to the grave nature of what Members of the Public are tasked to deliberate on the mere logical possibility of an empty eligible set is not enough to cause serious worry. For Gaus to seriously worry about this logical possibility it would need to be shown that these scenarios are the rule, not the exception.

In a similar vein, Riker's challenge thus far only highlights a logical possibility. It shows that different social choice mechanisms can, when applied to the optimal eligible set, select different moral rules. But if we are to hold the convergence solution to the same standards as Gaus holds his own justificatory framework to, then we must ask: what is the likelihood different social choice mechanisms applied to the optimal eligible choose the same moral rule? We are thus here concerned – and will be for the rest of the paper – with the plausibility of the convergence solution *when compared to* the plausibility of Gaus's own model of public reason, not the *absolute* plausibility of the convergence solution.[6] If it is likely that differing social choice mechanisms select the same moral rule then we have no cause for worry – the convergence solution succeeds, Nozick's objection is defeated, and Rawls's conclusion silenced. But if the chances are high that social choice mechanisms diverge on what moral rule is selected then we are indeed in trouble – in such a case Rawls's conclusion rings true. Riker certainly had views on this matter: "The moral and prudential standoff among [voting] methods would not in itself occasion difficulty for democratic theory if 'most of the time' most methods led to the same social choice from a given profile of individual values. But this is not the case" (Riker 1982: 235). Let us see if Riker is correct.

[6] I thank an editor at *Politics, Philosophy, & Economics* for drawing this distinction.

§4 The Convergence Solution and Real Public Reason

4.1 *Impartial Cultures and Divergence*.

In general there are five ways of examining with what frequency different social choice mechanisms converge on the same choice set (Felsenthal et al. 1993: 1-3; Felsenthal and Machover 1995: 144-147). First (*i*) we can examine analytically the mathematical properties of different social choice mechanisms and from there determine with what probability these mechanisms converge (probabilities not based on knowledge of actual behavior, preferences, associations, etc.) Second (*ii*) we can run computer simulations under different assumptions concerning the distribution of preference profiles and population sizes in society and from there estimate the frequency of convergence. Third (*iii*) we can use laboratory experiments to ask subjects to vote for the same set of candidates under different electoral schemes to see what level of convergence emerges. Fourth (*iv*) we can collect poll data on individual preferences and from there apply different social choice mechanisms to these preferences. And fifth (*v*) we can look at actual elections and apply different social choice mechanisms to the preferences given in the data sets to see if a different candidate would have been selected were a different social choice rule employed. In examining the robustness of the convergence solution in the face of Riker's challenge we focus on methods (*i*) and (*ii*), because these methods are the most developed in the existing literature and also allow for maximal flexibility when it comes to examining different sorts of preference distributions we might find in liberal societies.

Both analytic and simulation results are heavily dependent on the kinds of assumptions built into the model, particularly with what frequency we assume different preference orderings obtain at. The most common assumption concerning voter preferences is called the *impartial culture assumption*, as we shall see when examining several results that employ this assumption below. The impartial culture assumption says that for every individual in the population, there is an equal chance of that individual having one of the logically possible preference orderings given the feasible set of alternatives. As an example, suppose there are only three moral rules Members of the Public have preferences over: r_1, r_2, and r_3. From these three options there are six logically possible orderings Members may have (let $\succ$ stand for "is preferred to"): $(r_1 \succ r_2 \succ r_3)$, $(r_1 \succ r_3 \succ r_2)$, $(r_2 \succ r_1 \succ r_3)$, $(r_2 \succ r_3 \succ r_1)$, $(r_3 \succ r_1 \succ r_2)$, and $(r_3 \succ r_2 \succ r_1)$. The impartial culture assumption says that when we select any one Member of the Public at random that Member has an equal probability of adopting any one of these six orderings: there is a 1/6 chance Althea has ordering $(r_1 \succ r_2 \succ r_3)$, a 1/6 chance Althea has ordering $(r_1 \succ r_3 \succ r_2)$, etc.[7]

[7] Another assumption, the *impartial anonymous culture*, is also becoming common in the literature, though still less so than the more basic impartial culture assumption. The impartial anonymous culture says that each logically possible n-tuple of preferences for n voters is equally likely to occur. So, if we have a voting situation with n voters that creates x logically possible n-tuples of preferences, there is a 1/x chance the first logically possible n-tuple obtains, a 1/x chance the second logically possible n-tuple obtains, etc. We shall not explore impartial anonymous culture results in-depth because the implications are largely the same as those of the impartial culture, given our particular concerns. Namely (*i*) impartial anonymous cultures cause trouble (though slightly less trouble) for the convergence solution (see Gehrlein 1992: Table 1 for comparison), yet (*ii*) impartial anonymous cultures also cause problems for Gaus's model if such preferences obtain over the set of all proposed moral rules, in that Gaus's model will only yield a non-empty eligible set given an impartial anonymous culture if an n-tuple is selected satisfying the never worst domain restriction (discussed later in this section), which is very unlikely given a large population size. As such,

What happens when there is an impartial culture? When there is an impartial culture things look quite bleak for the convergence solution: in general, convergence rates of major social choice mechanisms – plurality voting, plurality runoff, the Hare method, approval voting, the Borda count, and the Coombs rule[8] – are all quite low. Holding the number of voters at 25, computer simulations show that plurality voting converges with Condorcet 42.6 percent of the time, plurality runoff 61.3 percent of the time, the Hare rule 77.9 percent of the time, approval voting 61.3 percent of the time, the Borda count 84.3 percent of the time, and the Coombs rule 81.1 percent of the time (Merrill 1988: Table 2.1). Such numbers do not constitute a particularly forceful response to Riker's challenge for the convergence solution.

Things get even worse when we increase the number of voters given an impartial culture: as population increases, rates of convergence decrease. With 999 voters for instance, plurality voting converges with the Condorcet method just 27 percent of the time, the Borda count 85 percent of the time, plurality runoff 45 percent of the time, the Hare method 75 percent of the time, and the Coombs rule 71 percent of the time (Nurmi 1992: Table 2). Again, not particularly inspiring numbers for the convergence solution in the face of Riker's challenge. Analytic results tend to back up these simulations, showing that (*i*) given an impartial culture convergence rates are quite low, and (*ii*) as population size increases these rates of convergence decrease (see, for example, Gehrlein 1985; Gehrlein and Lepelly 2000; Merlin et al. 2000).

---

extensive discussion of the impartial anonymous culture adds little to the main philosophical point of the paper.

[8] For those interested in the specifics of these social choice mechanisms as well as others, see Riker (1982: chp. 4) for an overview.

These results are all the more worrisome once we consider the likely number of "voters" we will be dealing with in Gaus's justificatory model. In terms of how this number is determined, "Members of the Public are… defined as the participants in the relevant rule-governed practice" (Gaus 2011: 268). If we are examining to what extent a rule governing circumstance C is justified, we take into account the preferences of all those citizens that participate in C, and thus their corresponding Member of the Public. That said, Gaus has a specific "focus on moral rules applying to a large-scale moral order in which we often confront others as strangers – what F.A. Hayek called 'the Great Society'" (Gaus 2011: 268). Given this focus on large-scale moral orders, we must assume that there are a very great many citizens engaged in the relevant rule-governed practice, and thus a very great many Members of the Public – certainly far more than the 999 voters in the above cited simulations that caused trouble for the convergence solution. Since the rates of convergence for all major social choice mechanisms tends to decrease as population increases given an impartial culture, and since we are theorizing about large-scale moral orders, these results are quite worrisome.

There are compelling reasons to discount such results, though. As discussed in §3.3, we are interested in holding the convergence solution to the same standards Gaus holds his own justificatory model to: for example, even though Riker's challenge points out the logical possibility of the convergence solution failing to answer Nozick's objection, such logical possibility gives us insufficient reason to reject the convergence solution because it is logically possible that the eligible set ends up empty on Gaus's own justificatory model. Similarly, we should not care about how impartial culture results bare on the convergence solution in the face of Riker's challenge because impartial cultures cause problems for Gaus as well, if we add one plausible assumption.

If we assume there is an impartial culture over the set of all proposed moral rules (before we narrow down to the eligible set and then optimal eligible set) and we assume – plausibly, I think –

that each Member of the Public at the very least places their last-ranked moral rule below blameless liberty, then it turns out that the eligible set ends up empty, meaning no moral rule is ultimately justified. This is displayed in Table 4. With three possible moral rules (r_1, r_2, and r_3), each logically possible ordering is represented at the same rate of occurrence, placing each moral rule below blameless liberty by at least two Members of the Public. Note, this does not hold for only three moral rules – the number of moral rules can be arbitrarily large, yet so long as there is an impartial culture and Members place their least preferred moral rule below blameless liberty, the eligible set will wind up empty. Because we are trying to work within and extend Gaus's justificatory model, we should treat any circumstances in which the eligible set winds up empty as also irrelevant for the convergence solution in the face of Riker's challenge – since Gaus's own model cannot handle such cases, it would be unfair to ask my model to. By implication, impartial culture results should not count as a refutation of the convergence solution.

Althea	Bertha	Cassidy	Dupree	Esau	Franklin
r_1	r_1	r_2	r_2	r_3	r_3
r_2	r_3	r_1	r_3	r_1	r_2
Blameless Liberty	Blameless Liberty	Blameless Liberty	Blameless Liberty	Blameless Liberty	Blameless Liberty
r_3	r_2	r_3	r_1	r_2	r_1

Table 4

What kinds of preferences does Gaus need Members of the Public to have over the set of all proposed moral rules in order for the eligible set to be non-empty? If we continue assuming that each Member places their least preferred rule below blameless liberty then for the eligible to be non-empty Gaus needs to employ a particular version Amartya Sen's *value restriction* assumption when it comes to the preferences of Members of the Public over the set of all proposed moral rules (Sen

1969/1982: 111). Specifically, Gaus must employ what is called by some the "never worst" version of value restriction (Regenwetter et al. 2006: 40). This never worst assumption is a bona fide domain restriction rather than a restriction on distributions of preferences like the impartial culture assumption. Never worst says that at least one moral rule will *never* be ranked worst by *all* Member of the Public.[9] As an example, if there are three proposed moral rules (r_1, r_2, and r_3), the never worst assumption says that one of these rules – either r_1, r_2, or r_3 – will not be ranked last by *each and every* Member. This means that when we select one Member of the Public at random, there is a zero percent chance she will have two of the six logically possible orderings. Then – and only then – will the eligible set be non-empty if we assume each Member places their least preferred moral rule below blameless liberty.

Clearly a domain restriction is quite an extreme restriction on permissible levels of diversity allowed to be present in society. It is especially harsh once we again remember that we are theorizing about Hayek's Great Society in which we are dealing with a very great number of people. As an example, if we took the relevant population to be the United States of America and assume for simplicity that only three moral rules are proposed, then in order for the eligible set to be non-empty *each and every* Member of the Public representing *each and every* citizen of the United States must adopt one of four of the six logically possible preference orderings that may appear. If merely *one* Member adopts *one* of the preference orderings that violate never worst, then the entire eligible set is empty, and nothing is publicly justified. Furthermore, the required domain restriction over the set of all

[9] Technically, never worst is a condition on triples: for all triples x, y, z never worst is satisfied so long as one of the alternatives is never ranked last in that triple by all voters. So with three voters and the triple x, y, z, the profiles {x, y, z}, {y, x, z}, and {z, x, y} satisfy never worst because one option, x, is not ranked last by any voter.

proposed moral rules becomes even more extreme if we assume that some Members of the Public place even more moral rules below blameless liberty. Indeed, if each Member places their two lowest-ranked rules below blameless liberty, then four of six logically possible preference orderings may not appear for each and every Member of our large-scale social order. If merely one Member adopts one of these four orderings then the eligible set winds up empty.

Since we are trying to hold the convergence solution to the same standards Gaus holds his own model to we can ask what happens to the convergence solution when there are similar restrictions on the diversity of rankings over optimal eligible set – note, we are not saying the optimal eligible set will necessarily be characterized by never worst; we are merely inquiring what happens when it is. Now there is some existing research that examines convergence rates of social choice mechanisms when domain restrictions are introduced. Lepelly et al. show that, for all constant scoring rules, if preferences satisfy the *single-peaked* domain restriction then, as population size increases, the Condorcet efficiency – the conditional probability that a rule will select the Condorcet winner, given that a Condorcet winner exists – of all such rules approaches one (Lepelly et al. 2000: Table 1).

Note, though, that a single-peaked domain restriction is equivalent to a never worst domain restriction. As a result, Lepelly et al. show that when we have a never worst restriction on preferences and the population gets large enough, *then there is perfect convergence on the Condorcet winner*. Gehrlein and Lepelly further show that, at least when it comes to the plurality rule, when preferences satisfy the *single-troughed* domain restriction (a close relative of the single-peaked assumption) then, again, there are high rates of convergence (Gehrlein and Lepelly 2009: 291). Thus, when we start looking at the kinds of preference distributions Gaus must assume over the set of all proposed moral rules for the eligible set to be non-empty we find that, if preferences over the optimal eligible set have a similar structure, then the convergence solution to Nozick's objection is quite compelling – all major social choice mechanisms end up selecting the same moral rule.

4.2 *Homogeneity and Convergence.*

In comparison to Gaus's justificatory model and the possibility of an empty eligible set, for the convergence solution to succeed it does not need anything nearly as extreme as a domain restriction. Indeed, once we start deviating from the impartial culture ever-so-slightly by introducing some homogeneity – defined broadly as the degree of consistency among voter preferences (more precise definitions of homogeneity are discussed below) (Gehrlein and Lepelly 2011: 42) – in terms of preference distributions over the optimal eligible set then the convergence of social choice mechanisms on the same choice set tends to increase as population size increases, which bodes well for us given our focus on Hayek's Great Society. Importantly, this introduction of homogeneity is not a radical restriction on the domain like never worst: all logically possible preference orderings may appear among Members of the Public at a non-zero probability. Gehrlein and Lepelly call this general trend of converging social choice mechanisms given increased homogeneity the *Efficiency Hypothesis*, which conjectures that "as voters have preferences on candidates that reflect increased levels of social homogeneity or group mutual coherence, voting rules should tend to show an increased level of Condorcet Efficiency" (Gehrlein and Lepelly 2011: 190).

As one example of the Efficiency Hypothesis, consider simulations run under the assumption that ten percent of the population endorses one particular preference ordering, whereas the other 90 percent of the population is an impartial culture. Under such an assumption the exact *opposite* relationship obtains between population size and convergence than that which obtains when there is an impartial culture: instead of decreasing, social choice mechanisms tend to *increase* in their rates of convergence as the number of voters increases. If you hold the number of voters at 25 and assume fifteen candidates then plurality runoff and the Hare method deviate 50 times, the Hare method and the Coombs method deviate 64 times, the Hare method and Nanson's method deviate 55 times, the Coombs method and Nanson's method deviate 49 times, the Copeland method and Nanson's method deviate 31 times, the Coombs method and the Copeland method deviate 43 times, and the Hare method and the Copeland method deviate 51 times. But, once we increase the number of voters from 25 to 999, then *there are zero deviations between all the above mechanisms* (Nurmi 1992: 482-485). That is, increasing the size of the population given a slightly homogenous culture leads to perfect convergence among varying social choice mechanisms. Given our focus on large-scale moral orders, such results are encouraging for the convergence solution.

Importantly, it should be noted that if preferences over the set of all proposed moral rules satisfies the preference distribution used to obtain the above results then the eligible set ends up empty. For though we are placing some restrictions on distributions of preferences when we introduce homogeneity we do not place any restrictions on the domain of preferences: each logically possible ordering occurs at a non-zero probability. Thus, if we continue assuming that each Member of the Public places their least preferred moral rule below blameless liberty then if preferences over the set of all moral rules satisfy the above distribution each moral rule is vetoed by some Member. But as can be seen, such a distribution, when over the optimal eligible set rather than the set of all proposed moral rules, poses no problem for the convergence solution. The reasons why should be obvious: when it comes to preferences over the set of all proposed moral rules, every Member is effectively given veto power, such that one Member placing a moral rule r_n below blameless liberty makes r_n no longer a candidate for public justification. But with the convergence solution no such veto power is given: all logically possible orderings can obtain yet, still, there will more likely than not be perfect convergence once we introduce some homogeneity and let the population size get large enough.

The above results test one particular instance of homogeneity and shows high rates of convergence as population size increases. In fact, there are more general tests we can run: we can define homogeneity in a broad way that is able to capture preference distributions such as the one employed above along with other instances of homogeneity as well. From there, we can use simulations to test rates of convergence as these levels of homogeneity increase, as well as an increase in the number of voters. Generally speaking there are two different measures of homogeneity widely appealed to in the literature, one relying on Abrams's (1976) definition, which tracks the degree of agreement on particular preference orderings, and another relying on Kendall's coefficient of concordance (Fishburn 1973), which tracks the degree of deviation from an impartial culture towards what is called a *purely arbitrary culture*, which is defined as a population with perfect agreement on one preference ordering.

Gehrlein (1987; 1995) and Gehrlein and Lepelly (2011: 191-194) show that, once we start introducing homogeneity of both kinds defined above then (*i*) several major social choice mechanisms begin increasing in convergence as the level of homogeneity increases, and (*ii*) this convergence increases still as population size increases. Indeed, once the population size gets large enough then, given the introduction of some homogeneity, we should see perfect convergence of social choice mechanisms. The rules for which these relationships hold are plurality rule, the Borda count, plurality runoff, and negative plurality runoff. The only exception here is negative plurality rule, which *decreases* in Condorcet efficiency as homogeneity increases. Here, one might think that we have a quite serious problem: after all, the convergence solution does not require *most* social choice mechanisms of our Members of the Public all select the same rule; it requires *all* social choice mechanisms do. If at least one Member most prefers negative plurality then, even when we have slight homogeneity, things look quite bleak for the convergence solution.

In response, the performance of negative plurality rule does not cut against the convergence solution, for reasons given in §3.1. There we said that ties are not permissible – all social choice mechanisms must be decisive, in that they must select one and only one winner. But, as Lepelly et al. (2000: 190) note, the reason why negative plurality performs so poorly as social homogeneity increases in terms of its Condorcet efficiency is because it selects more than one candidate as the winner, yet there is only one Condorcet winner. That is, negative plurality *does* select the Condorcet winner as social homogeneity increases, but it selects other winners as well. Since Condorcet efficiency measures when a rule selects the Condorcet winner *and only* the Condorcet winner, this accounts for negative plurality's low efficiency when social homogeneity increases. Because we stipulate negative plurality cannot behave like this given the formal features of Gaus's justificatory apparatus, we do not have to worry about such results indicting the convergence solution.

The results canvassed in this section show that, even though impartial culture assumptions create a rather bleak outlook for the convergence solution, slight deviations from the impartial culture should inspire optimism – as Gehrlein and Lepelly's Efficiency Hypothesis states, once we begin introducing homogeneity we should start seeing convergence of all major social choice mechanisms on the Condorcet winner, especially as population size grows. But more crucially than that, such homogeneity is *not* sufficient to yield a non-empty eligible set when characterizing preferences over the set of all moral rules – under plausible assumptions concerning when Members place moral rules below blameless liberty, all the above distributions result in no moral rule being justified. This is essential, for our goal in amending Gaus's theory is to show that our annexation at least succeeds under relatively similar standards he himself holds his model to.

4.3 *Preference Patterns and Robustness*.

It has been argued (*i*) that the kinds of domain restrictions Gaus requires over the set of all proposed moral rules for the eligible set to be non-empty, when applied to the optimal eligible set, bode well for the convergence solution; and (*ii*) that less severe and more plausible diversity restrictions over the optimal eligible set that are sufficient for the convergence solution to perform well, when applied to the set of all proposed moral rules, lead to trouble for Gaus's basic justificatory model. In terms of robustness, then, the convergence solution outperforms Gaus's model: it succeeds under those very same diversity conditions Gaus's model succeeds under, and then succeeds under some diversity conditions Gaus's model does not succeed under. Thus, if one has no complaint against the robustness of Gaus's basic model, then one can have no complaint against my extension of it – his model will yield instances of nothing be justified more frequently than mine does.

But note, the above claim is true and this general line of argument only compelling given a certain assumption that so far has gone unstated: that all kinds of diversity in terms of preference distributions are equally likely to occur. If one could either argue (*iii*) that never worst domain restrictions are likely to characterize preferences over the set of all moral rules, or (*iv*) that impartial cultures rather than homogeneity assumptions are likely to characterize the optimal eligible set, then one could effectively rebut the argumentative strategy §4 of this paper has employed. I cannot think of an argument in defense of (*iv*), but I would like to end by entertaining an argument in defense of (*iii*).[10]

[10] I thank an anonymous reviewer at *Politics, Philosophy, & Economics* for bringing this argument to my attention.

One reason for thinking that preferences over the set of all proposed moral rules are likely characterized by never worst value restriction is due to a sort of "early agreement" among Members of the Public – some options r_n are so unattractive to all Members that other options, say r_m, will never be ranked worst because those deeply unattractive r_n will always be ranked below r_m. As a concrete example of what this might look like, r_n might be a rule that makes everyone worse off compared to r_m and, as a result, it is plausible to think that r_m will never be ranked last among all Members of the Public, for r_n will always be ranked below it. We thus have an instance of never worst. Note, this line of reasoning could not be used to argue that the optimal eligible set is likely characterized by never worst – for here, if there really was an r_n so unattractive that all Members preferred r_m to it, then such an r_n could not be in the optimal eligible set in the first place, for by definition the optimal eligible set contains no Pareto dominated elements. What allows one to plausibly argue that the set of all proposed moral rules likely contains perverse rules of the kind described above is due to the fact that the set of all proposed moral rules has undergone no refinement – once pruning has begun (as in the case of the optimal eligible set), then rules of the kind described above will be eliminated, and arguments of this kind off-limits.

Now whether the set of all proposed moral rules likely looks like this depends on how the set of all moral rules Members have preferences over is determined. If we were simply considering the set of all *possible* moral rules then perhaps it is likely there will always be some options so unattractive or perverse that in comparison other options will never be ranked below them. Gaus's justificatory model is more modest than this, however. On Gaus's model, each Member of the Public may, but is not required to, propose any moral rule she pleases (Gaus 2011: 293-294). We thus do not begin by considering all possible moral rules, but only those that are proposed by Members. Since this is true, it is hard to see why there would *frequently* be rules r_n and r_m such that the relationship posited above obtains – where r_n is so perverse a rule that *everyone* prefers r_m to r_n, making it such that r_m will never be worst. For if r_m is on the table, then why would anyone ever propose r_n, given how r_n is defined? Now of course, it could be that r_n was proposed first followed by r_m, making the newly proposed r_m never worst. But this mere possibility of ordering effects is not sufficient to show that the set of all eligible rules is *likely* to be characterized by never worst – only that it indeed *can* be characterized by never worst. As a result, the current objection fails.

§5 Conclusion

One of the key features of Gaus's public reason liberalism project is a deep commitment to realism. This is why Gaus does not idealize away the diversity present in liberal societies when he models Members of the Public, in contrast to Rawls's original position and those who narrowly follow Rawls's path. The goal of public reason liberalism and the project of seeking a justified social morality is to see, when we demand that others obey moral rules, whether we actually have authority to make such claims, or whether we are just "pushing people around" (Gaus 2011: 16). To make sure we take seriously the possibility that we *are* just pushing people around we do not idealize away the diversity we find around us, leading us to the indeterminacy problem.

Note that, given the nature of the question Gaus asks – whether you and me, here and now, are justified when we make authoritative claims on one another – the style of theorizing must change from what is traditionally done in philosophy. Mere logical possibility is no longer sufficient to defeat a theory (for, again, the theory is for you and me, here and now); we need empirical likelihood to do that. Gaus argues that his theory meets such a standard – that the optimal eligible set will likely not be empty, meaning that it is possible for us to have a publicly justified social morality. I have argued that my extension of Gaus's theory also meets such a standard – that social choice mechanisms will likely converge on the same moral rule, solving the indeterminacy problem in a manner that secures the moral status of the resulting equilibrium, as required by criterion (c). Like Gaus's basic justificatory model, we can of course define cases for which the convergence solution fails. But for the kinds of problems you and I are most interested in theorizing about, the convergence solution shows how we can solve the indeterminacy problem in a manner that treats all as free and equal. Indeed, recall that Rawls's conclusion questions whether it turns out that, *for us*, there exists a reasonable and workable conception of justice at all. The convergence solution, annexed to Gaus's broader theory, shows that there can be such a conception.

Works Cited

Abrams, R. 1976. "The Voter's Paradox and the Homogeneity of Individual Preference Orders." *Public Choice* 26: 19-27.

Arrow, Kenneth. 1951/2012. *Social Choice and Individual Values*. New Haven: Yale University Press.

Bowles, Samuel and Herbert Gintis. 2011. *A Cooperative Species: Human Reciprocity and its Evolution*. Princeton: Princeton University Press.

D'Agostino, Fred. 2013. "The Orders of Public Reason." *Analytic Philosophy* 54: 129-155.

Dowding, Keith and Martin van Hees. 2008. "In Praise of Manipulation." *British Journal of Political Science* 38: 1-15.

Felsenthal, Dan S. and Moshé Machover. 1995. "Who Ought to be Elected and Who is Actually Elected? An Empirical Investigation of 92 Elections Under Three Procedures." *Electoral Studies* 14: 143-169.

Felsenthal, Dan S., Zeev Maoz, and Amnon Rapport. 1993. "An Empirical Evaluation of Six

Voting Procedures: Do They Really Make Any Difference?" *British Journal of Political Science* 23: 1-27.

Fishburn, Peter C. 1973. "Voter Concordance, Simple Majority, and Group Decision Methods." *Behavioral Science* 18: 364-376.

Gaus, Gerald. 2011. *The Order of Public Reason.* Cambridge: Cambridge University Press.

Gehrlein, William V. 1985. "Condorcet Efficiency of Constant Scoring Rules for Large Electorates." *Economics Letters* 19: 13-15.

Gehrlein, William V. 1987. "The Impact of Social Homogeneity on the Condorcet Efficiency of Weighted Scoring Rules." *Social Science Research* 16: 361-369.

Gehrlein, William V. 1992. "Condorcet Efficiency of Simple Voting Rules for Large Electorates." *Economics Letters* 40: 61-66.

Gehrlein, William V. 1995. "Condorcet Efficiency and Social Homogeneity." In *Social Choice, Welfare, and Ethics: Proceedings of the Eight International Symposium on Economic Theory and Econometrics*, edited by William A. Barn: 127-143. Cambridge: Cambridge University Press.

Gehrlein, William V. and Dominique Lepelly. 2000. "The Probability That All Weighted Scoring Rules Select the Same Winner." *Economics Letters* 66: 191-197.

Gehrlein, William V. and Dominique Lepelly. 2009. "The Unexpected Behavior of Plurality Rule." *Theory and Decision* 67: 267-293.

Gehrlein, William V. and Dominique Lepelly. 2011. *Voting Paradoxes and Group Coherence: The Condorcet Efficiency of Voting Rules*. Heidelberg: Springer.

Gibbard, Allan. 1973. "Manipulation of Voting Schemes: a General Result." *Econometrica* 41: 587-601.

Lepelly, Dominique, Patrick Pierron, and Fabrice Valognes. 2000. "Scoring Rules, Condorcet Efficiency, and Social Homogeneity." *Theory and Decision* 49: 175-196.

List, Christian and Philip Pettit. 2002. "Aggregating Sets of Judgments: An Impossibility Result." *Economics and Philosophy* 18: 89-110.

Merlin, V., M. Tataru, and F. Valognes. 2000. "On the Probability that all Decision Rules Select the Same Winner." *Journal of Mathematical Economics* 33: 183-207.

Merrill, Samuel. 1988. *Making Multicandidate Elections More Democratic*. Princeton: Princeton University Press.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.

Nurmi, Hannu. 1992. "An Assessment of Voting System Simulations." *Public Choice* 73: 459-487.

Rawls, John. 1971. *A Theory of Justice*. Cambridge: Harvard University Press.

Rawls, John. 1980/1999. "Kantian Constructivism in Moral Theory." In *Collected Papers*: 303-358. Cambridge: Harvard University Press.

Regenwetter, Michel, Bernard Grofman, A.A.J. Marley, Iliam M. Tsetlin. 2006. *Behavioral Social Choice*. Cambridge: Cambridge University Press.

Riker, William. 1982. *Liberalism Against Populism.* Long Grove: Waveland Press.

Satterthwaite, Mark A. 1975. "Strategy-Proofness and Arrow's Conditions." *Journal of Economic Theory* 10: 187-217.

Sen, Amartya. 1969/1982. "A Possibility Theorem on Majority Decisions." In *Choice, Welfare and Measurement*: 109-117. Cambridge: Harvard University Press.