Rule-Consequentialism's Assumptions

KEVIN P. TOBIA

Yale University

Abstract: Rule-Consequentialism faces "the problem of partial acceptance": How should the ideal code be selected given the possibility that its rules may not be universally accepted? A new contender, "Calculated Rates" Rule-Consequentialism claims to solve this problem. However, I argue that Calculated Rates merely relocates the partial acceptance question. Nevertheless, there is a significant lesson from this failure of Calculated Rates. Rule-Consequentialism's problem of partial acceptance is more helpfully understood as an instance of the broader problem of selecting the ideal code given various assumptions—assumptions about who will accept and comply with the rules, but also about how the rules will be taught and enforced, and how similar the future will be. Previous rich discussions about partial acceptance provide a taxonomy and groundwork for formulating the best version of Rule-Consequentialism.

Keywords: Rule-Consequentialism, Rule Utilitarianism, Fixed Rate, Variable Rate, Optimum Rate, Top Rate, Maximizing Expectation Rate, Calculated Rate

I. RULE-CONSEQUENTIALISM AND THE PROBLEM OF PARTIAL ACCEPTANCE

Rule-Consequentialism's modern "partial acceptance" debate begins with Brad Hooker's articulation of Fixed Rate Rule-Consequentialism:

> An act is wrong if and only if it is forbidden by the code of rules whose internalization by the overwhelming majority of everyone everywhere in each new generation has maximum expected value in terms of well-being . . . The calculation of a code's expected value includes all costs of getting the code internalized.[1]

What exactly is the "overwhelming majority of everyone everywhere"? Ninety-percent, answers Hooker.[2]

This ninety-percent answer is a "Fixed Rate" response to Rule-Consequentialism's "problem of partial acceptance": How should the ideal code be selected given the possibility that its rules may not be universally accepted?[3]

But a Fixed Rate response is not the only answer to the problem of partial acceptance. Michael Ridge offers a "Variable Rate" response; we should select the code with the greatest expected value across all possible levels of acceptance.[4] Holly Smith defends an "Optimum Rate" answer; we should select the code with the greatest expected value at some level of acceptance.[5] Another option is a "Maximizing Expectation Rate" theory; the ideal code is that whose expected value is greatest as an average across all possible acceptance levels, weighted by the probability of each level obtaining.[6] Teemu Toppinen argues for a "Top Rate" response: choose the code whose acceptance by one-hundred percent has greatest expected value.[7] For a summary of these views, see Figure 1.

[1] Brad Hooker, *Ideal Code, Real World* (Oxford, 2000), p. 32.

[2] Hooker, *Ideal Code, Real World*, p. 83; see also Richard Brandt, *Morality, Utilitarianism, and Rights* (Cambridge, 1992).

[3] A referee asks a helpful and neglected question: why exactly is the fact that no code will secure universal acceptance a problem? Once we identify the "ideal code," why would partial acceptance challenge its status as the ideal code? To see the relevance of acceptance, imagine that we identify the seemingly ideal moral code, but estimates reveal that it will be accepted by zero percent of the population. It is hard to see how any plausible conception of Rule-Consequentialism—concerned with the value of consequences—could endorse this as the ideal code.

[4] Michael Ridge, 'Introducing Variable Rate Rule Utilitarianism', *Philosophical Quarterly* 56 (20016), pp. 242–253.

[5] Holly Smith, 'Measuring the Consequences of Rules', *Utilitas* 22 (2010), pp. 413-433.

[6] Kevin Tobia, 'Rule Consequentialism and the Problem of Partial Acceptance', *Ethical Theory and Moral Practice* 16 (2013), pp. 643-652.

[7] Teemu Toppinen, 'Rule Consequentialism (and Kantian Contractualism) at Top Rates', *Philosophical Quarterly* 66 (2016), pp. 122-135.

Theory	The Ideal Code Is . . .	Example
Fixed Rate	That whose acceptance by ninety percent has the greatest expected value	Hooker 2000
Top Rate	That whose acceptance by one-hundred percent has the greatest expected value	Toppinen 2016
Variable Rate	That whose expected value is greatest on average across all levels of acceptance	Ridge 2009
Optimum Rate	That whose optimum acceptance level has the greatest expected value	Smith 2010
Maximizing Expectation Rate	That whose expected value is greatest on average across all levels of acceptance weighted by the probability of acceptance at each level	Tobia 2013

Figure 1. Five responses to Rule-Consequentialism's "Problem of Partial Acceptance": How should the ideal code be selected given that its rules may not be universally accepted?

These theories recommend different possible codes as the ideal code. Which code each theory identifies depends on various facts about the world and what possible alternative codes exist. For instance, consider Figure 2 below. In this toy example, there are only two possible codes. The Y-axis depicts expected value[8] and the X-axis depicts possible acceptance levels. On all five theories, Code 1 is identified as the ideal code. Compared to Code 2, Code 1 promises greater expected value at ninety percent acceptance (Fixed Rate), at one-hundred percent acceptance (Top Rate), on average across all acceptance levels (Variable Rate), at its optimum acceptance level of eighty percent (Optimum Rate), and on any average weighted by the probability of acceptance at each level (Maximizing Expectation Rate).[9]

[8] Throughout the article I use "value" rather than "welfare" or "utility" to follow Hooker, *Ideal Code, Real World*.
[9] There is no need to specify the likelihood distribution in this example, since Code 1 strictly dominates Code 2. Given any distribution of acceptance levels, Code 1 has greater expected value than Code 2.

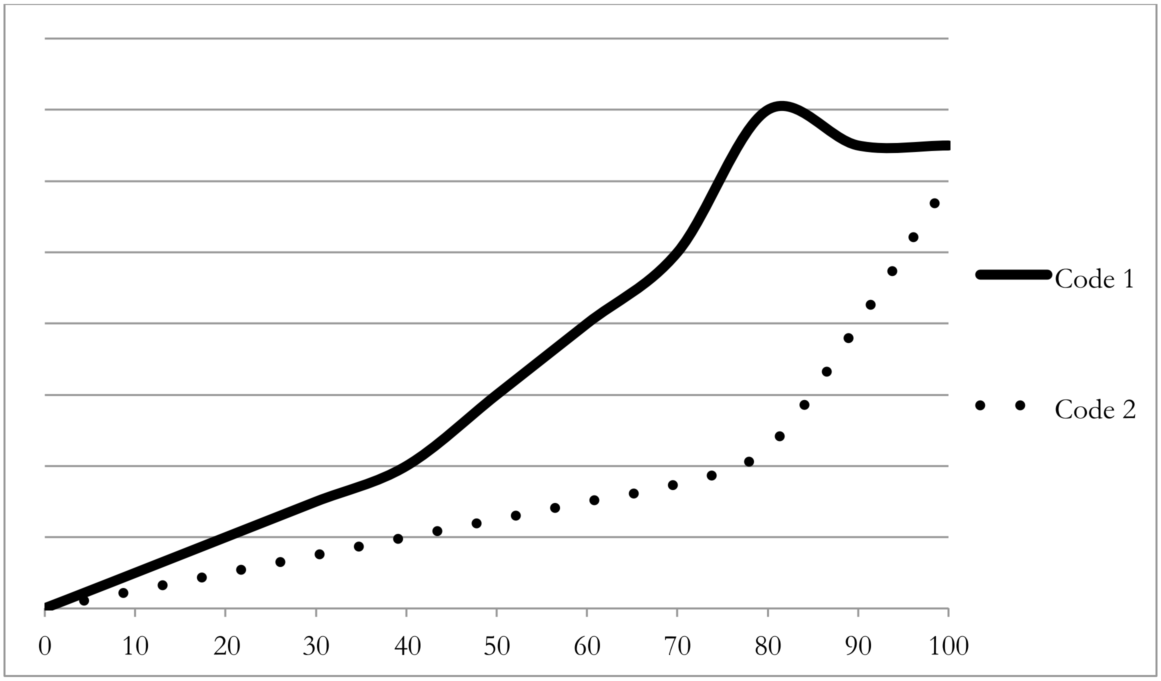

Figure 2. Expected value of two codes across possible acceptance levels.

In the real world, there are more than two possible codes. And their expected value functions at various acceptance levels are more complex. For some codes, it is reasonable to expect great value at higher levels of acceptance and low value at lower levels; for other codes the inverse expectation is reasonable. For a more complex example, see Figure 3 below.

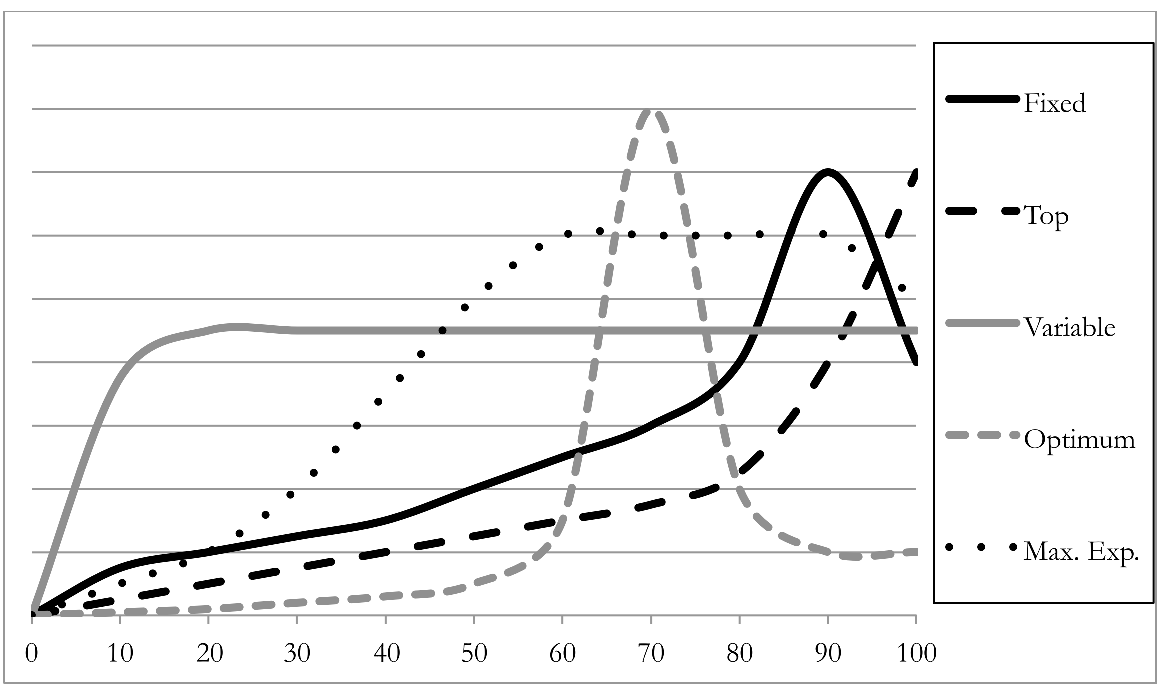

Figure 3. Expected value of five codes across possible acceptance levels. Each code is labeled by the theory under which it is the ideal code (assuming these five codes were exhaustive).[10]

This figure depicts a world with five possible codes. In this example, intended to illustrate the competing theories' differences, the codes are such that each of the five theories identifies a different code as the ideal one. Assume that acceptance levels are normally distributed around the seventy-five percent acceptance level. The Fixed Rate code is that with the greatest expected value at ninety percent acceptance; the Top Rate code is that with the greatest expected value at one-hundred percent acceptance; the Variable Rate code is that with the greatest expected value on average across all levels; the Optimum Rate code is that with the greatest expected value at its optimum (in this case, seventy percent); and the Maximizing Expectation Rate code is that with the greatest expected value when averaging across all possible acceptance levels, weighted by their likelihood of obtaining. Figure 3 indicates the significant differences between these competing versions of Rule-Consequentialism.

Which of these formulations of Rule-Consequentialism is best? This is the subject of much debate. Each of these theories has been critiqued individually, and all have also been critiqued collectively.

Fixed Rate theory is maligned as arbitrary;[11] why focus on ninety percent rather than eight percent or ninety-five percent? Moreover, it focuses on just one acceptance level. This leads to implausible consequences when a code performs well at ninety percent but poorly at many other levels.[12]

Top Rate theory is the newest addition to the debate, but it is subject to many of the same critiques as is Fixed Rate theory. It is less arbitrary than (ninety percent) Fixed Rate theory, but by focusing on just one acceptance level (one-hundred percent), Top Rate theory still risks identifying "ideal codes" that perform well at that single rate, but poorly at many others. This is particularly troublesome in so far as one-hundred percent acceptance is unlikely to obtain.

Variable Rate theory is also arbitrary; why does it recommend the *mean* across acceptance levels and not the median or mode?[13] More importantly, it is prone to being skewed by anomaly.[14]

Optimum Rate theory does not consider whether a given code is at all *likely* to achieve those optimal acceptance rates. A risky code with an improbable optimum is not obviously the ideal one.[15]

Maximizing Expectation Rate theory has not been subjected to any individual critiques.[16] A tempting criticism is that its selection procedure is more complicated than

[10] And that the probability of an acceptance level obtaining is normally distributed over seventy-five percent. This assumption is required to distinguish the Maximizing Expectation Rate code from the Variable Rate ideal code.

[11] Ridge, 'Introducing Variable Rate Rule Utilitarianism'.

[12] Tobia, 'Rule Consequentialism and the Problem of Partial Acceptance'.

[13] Tobia, 'Rule Consequentialism and the Problem of Partial Acceptance', p. 645.

[14] Brad Hooker and Guy Fletcher, 'Variable vs. Fixed Rate Rule Utilitarianism', *Philosophical Quarterly* 58 (2008), pp. 344-352.

[15] Tobia, 'Rule Consequentialism and the Problem of Partial Acceptance'; Timothy D. Miller, 'Solving Rule-Consequentialism's Acceptance Rate Problem', *Utilitas* 28 (2016), pp. 41-53.

that of other theories. This is not to say the theory is not simple; the maximization of expected value is a straightforward approach, and one that is less arbitrary than focusing on fixed rates. But the complexity might raise implementation problems.

These theories have also been critiqued collectively. For instance, Doug Portmore describes Maximizing Expectation Rate as the "best version . . . perhaps," but he also levels critiques that apply to all versions.[17] Holly Smith argues that indeterminacy of acceptance and compliance presents a challenge to all theories.[18] I turn now to a most recent competitor to all of these views, one that claims to "solve" the problem of partial acceptance.

II. CALCULATED RATES RULE-CONSEQUENTIALISM (CRCC)

Timothy Miller offers a new competitor theory to those discussed in the previous section.[19] This is "Calculated Rates" Rule-Consequentialism (CRRC). Miller claims that CRRC "has considerable advantages over [all these other] formulations" and "no disadvantages that those theories do not share." Miller's view is that Rule-Consequentialism should simply *solve* for the appropriate acceptance levels, rather than treat acceptance rate as an uncertain variable.

CRCC holds that "[a]n *ideal code* is a *code* included in an *expanded code* whose value is at least as high as the value of any alternative *expanded code*."[20] This definition includes Miller's original terminology. "*Code*" refers to a complete set of moral rules, and "*code + stipulated methods of teaching and enforcement*" is an "*expanded code*."[21] In other words, for any given code we can consider various methods of teaching and enforcement. For each of these methods, we calculate the expected acceptance and compliance of various rules within the code.

The fundamental problem with CRCC is that, although it seems to "solve" the partial acceptance problem, it merely masks it. Imagine a world with four possible codes, each of which can be taught and enforced by four different methods. Figure 4 depicts a matrix indicating the welfare or "value" of each of these Code x Method pairs.

---

[16] See Miller, 'Solving Rule-Consequentialism's Acceptance Rate Problem' for a critique that extends to all theories.

[17] Douglas W. Portmore, 'Parfit on Reasons and Rule Consequentialism,' in *Reading Parfit: On What Matters* (Simon Kirchin ed., 2017), p. 19.

[18] Smith, 'Measuring the Consequences of Rules.' But see Shang Long Yeo, 'Measuring the Consequences of Rules: A Reply to Smith', *Utilitas* 29 (2017): 125-131. For the distinction between acceptance and compliance, see Timothy Mulgan, *Future People: A Moderate Consequentialist Account of Our Obligations to Future Generations* (Oxford University Press, 2006), p. 138.

[19] Miller, 'Solving Rule-Consequentialism's Acceptance Rate Problem.'

[20] Miller, 'Solving Rule-Consequentialism's Acceptance Rate Problem', p. 50.

[21] Miller, 'Solving Rule-Consequentialism's Acceptance Rate Problem', p. 50.

	Code 1	**Code 2**	**Code 3**	**Code 4**
Method 1	1	2	8	4
Method 2	5	2	6	4
Method 3	5	10	2	4
Method 4	9	2	2	4

Figure 4. Expected Value Matrix: Methods of Training and Enforcement by Code

CRRC endorses choosing the code "whose value is at least as high as the value of any alternative expanded code." Now, which cell contains the code "whose value is at least as high" as the value of any alternative? Seemingly Code 2 (and Method 3).

But a crucial question remains: How are these matrix values calculated? Answering this question returns us to familiar territory of the problem of partial acceptance. Should the "values" that fill this table represent a Code x Method's Optimum at some level of acceptance, or the (Variable Rate) average across all possible levels of acceptance, or the reasoning of one of the alternative theories? In other words is "10", the value of Method 2 and Code 3, an indication of what those methods will arrive at in the best case acceptance scenario (Optimum), if they achieve one-hundred percent acceptance (Top) or ninety percent acceptance (Fixed), on average (Variable) or on a likelihood weighted average (Maximizing Expectation)?

Miller's solution is to simply "solve" for the acceptance rate that will obtain from a method. But a single method does not guarantee an acceptance rate, either for a specific code or across all codes. For example, "Method 3" cannot be understood as purchasing say a ninety percent acceptance rate across all different moral codes. Different codes have different inculcation costs. And within a single pair, say Method 3 and Code 2, we need to say more about how the acceptance rate is calculated.

I interpret Miller's suggestion as most plausibly endorsing Optimum or Maximizing Expectation reasoning to answer this question. In other words, teaching Code 2 with Method 3 will in the best case scenario (Optimum) achieve the level of acceptance providing 10 value, or the weighted average of acceptance levels (Maximizing Expectation) provides 10 value. But even such an interpretation requires a defense; why, on Miller's view, do we prefer a certain method of "solving for" acceptance?

More importantly, this reveals that CRCC is not a true competitor to these other theories. It is ultimately dependent on one of these other views, arguing that we should choose the Code x Method combination that offers the best value given an acceptance rate calculation—value ultimately calculated in terms of Variable or Optimum or Maximizing Expectation or some other theory.

This first objection concerns a problem in calculating acceptance levels given stipulated methods of training and enforcement. A second line of objection also problematizes CRCC. Is it really plausible that we can stipulate what methods of training and enforcement will obtain? In the real world, this is a question influenced by various political processes and complex social norm interactions. The far more plausible assumption is that we can estimate the likelihood of certain methods obtaining, given our intention for them to obtain.

Similar debates to the partial acceptance debates arise here: the various positions include endorsing the code that does best on average across various methods, the code that does the best at one possible method, or the code that does the best as a weighted average. Consider Figure 5, which depicts a Code x Method expected value matrix. This figure also reflects the more realistic assumption that attempting to teach and enforce some method may lead to slightly different methods actually obtaining—a slight difference with significant consequences.

	Code 1	**Code 2**	**Code 3**	**Code 4**
Method 1 [0.6]	1	2	8	4
Method 2 [0.2]	5	2	6	4
Method 3 [0.1]	5	10	2	4
Method 4 [0.1]	9	2	2	4

Figure 5. Expected Value Matrix: Methods of Training and Enforcement by Code, brackets indicate probability of the method actually obtaining

Let us set aside the first objection. Ignore the question of how these matrix values are calculated and take them as given. Consider just the second objection: The fact that the outcome of methods of training and enforcement are at best probabilistic outcomes requires us to make hard choices in ideal code choice. Figure 5 suggests different predictions of the same theoretical approaches to the problem of partial acceptance. For example, in Figure 5's matrix "variable methods theory" would select Code 1; "optimum methods theory" would select Code 2; "Maximum Expectation methods theory" would selects Code 3. When we cannot guarantee that an intended method will obtain, then "Calculated Rates" theory by itself does not give a clear prediction about the ideal code.

It is worth reiterating that this problem does not require large differences among possible methods of training and enforcement. We might understand Figure 5 as representing the probabilities obtaining from the intention or expected effort to implement some single method, say Method 1. We expect that intended implementation to actually result in one of four (perhaps subtly different) outcomes: Method 1, 2, 3, or 4. Those method outcomes interact differently with different codes, generating this second problem for Calculated Rates.

Together, these two critiques reveal that CRRC raises as many problems as it claims to solve. Attempting to "solve" for acceptance rates by endorsing some particular methods still raises questions about what kind of acceptance value should be calculated and how to respond to the partial success of methods of training and enforcement.

III. RULE-CONSEQUENTIALISM'S BROADER PROBLEM

The previous section identifies two problems with the Calculated Rates suggestion. Calculated Rates simply relocates the problem of partial acceptance. The guidance to select the code in an expanded code of greatest value is indeterminate; we need to know how the values are calculated, and answering that implicates the core problem of the partial acceptance debate. Second, it assumes that we can stipulate precise levels of training and enforcement. Insofar as there is any probabilistic variability in method outcome, we face similar problems to those raised by the problem of partial acceptance.

These critiques reveal the problem of partial acceptance as one instance of a broader problem for Rule-Consequentialism: Rules are chosen *given* various assumptions. The partial acceptance debate has focused on questions about how many people will accept the ideal code, but we should also consider questions about the realizability of methods of training and enforcement. Other factors are also relevant. For instance, we should consider the similarity of our current context to future contexts; a code that realizes great value today may not be as valuable in a future sufficiently dissimilar to ours. Alongside the "problem of partial acceptance," we have at least the "problem of training and enforcement" and the "problem of uncertain futures."

Rule-Consequentialism's broader problem is the problem of its factual stipulations or assumptions made under uncertainty. The core of Rule-Consequentialist analysis is evaluation of different codes' outcomes. The value of those outcomes depends crucially on the relationship between a code's rules and various factors (e.g. partial acceptance, partial compliance, other background factors like training and enforcement, dis/similarity of the future).

Although the partial acceptance debate is somewhat narrow, the various approaches outlined by that debate are useful to these other questions. Those theories provide a basic taxonomy of useful approaches and a groundwork for developing the best formulation of Rule-Consequentialism. See Figure 6 below for a summary of these approaches.

Fixed theories offer a brightline rule. For example, the ideal code is the code that does best when accepted by ninety percent, is complied with by the same, in ninety percent effective conditions (e.g. training and enforcement) similar to our current one.

Top theories also offer a clear brightline rule, for ideal conditions. For example, the ideal code is that at one-hundred percent acceptance, complied with by same, in conditions like ours.

Variable theories offer an averaged view. For example the ideal code is the code that does best on average over possible levels of acceptance and compliance, in various conditions.

Optimum theory selects the best possible code. For example, the ideal code is the code that does best for some acceptance and compliance rate, in some condition (perhaps not so similar to our current one).

Theory	**Given uncertainty of X (acceptance, compliance, training & enforcement, similarity of future, etc.), the Ideal Code is . . .**
Fixed	That whose X at the ninety percent level has the greatest expected value
Top	That whose X at the one-hundred percent has the greatest expected value
Variable	That whose expected value is greatest as an average across all levels of X
Optimum	That whose optimum X level has the greatest expected value
Maximin	That whose minimum X level has the greatest expected value
Maximizing Expectation	That whose expected value is greatest as an average across all levels of X weighted by the probability of each level obtaining

Figure 6. Theoretical approaches to the Problem of Rule-Consequentialism's Assumptions

A previously undefended option is a "Maximin theory," selecting the code that avoids the worst possible consequences. For example, the ideal code is the code that does least badly across all various conditions.

Maximizing Expectation theory selects the ideal code as the code that does best in an expected value calculation, taking acceptance and compliance rates and background conditions as probability weighted inputs. Further refinements are possible. For instance, "Plausible Expectation" theory might select the code that maximizes expected value, only over plausible outcomes levels (i.e. excluded from analysis improbable values). "Anomaly Avoidance" theory might first exclude all codes that have a probability of leading to unusually bad outcomes under certain conditions, and then apply Maximizing Expectation theory. There are countless other possibilities.

To see the relevance of these approaches across various Rule-Consequentialist debates, consider as an extended example the problem of uncertain futures. How should we identify the ideal code given that the conditions of the future might not mirror those of the present? One approach might be to only consider futures that are "close" but not identical to ours, and choose the code that does best in those contexts. This is essentially a "Fixed" style suggestion. Given the problematic variability in some parameter (e.g.

partial acceptance, future similarity), we should just pick a clear and simple cutoff and choose the code that performs best at the value (e.g. ninety percent acceptance, ninety percent similarity).

But perhaps what we should really consider is the code whose acceptance leads to the greatest value in the future *most* similar to our. This is a "Top" style suggestion. Given the problematic variability in some parameter, we should just choose the code that performs best at the value of the parameter that would in itself seem ideal (e.g. one-hundred percent acceptance, one-hundred percent future similarity).

Or we could look to the code whose expected value is greatest as an average across all possible futures (Variable). Or we could choose the code whose optimum in any future has greatest expected value (Optimum). Here, as elsewhere, the Maximizing Expectation style solution is highly plausible: the ideal code is the one whose expected value is greatest as an average across all future possibilities, weighted by the probability of each level obtaining.

We might wonder how broad a problem this presents for Rule-Consequentialism. For one, some of these issues might be considered together. For example, perhaps problems of partial acceptance and partial compliance can be dealt with swiftly using the same general Rule-Consequentialist assumption and endorsing one of the theoretical approaches (e.g. fixed or top or variable).

While there might be some reason to adopt a similar approach about certain issues (e.g. adopting the same theoretical approach with respect to both acceptance and compliance), there may be reasons to favor different approaches to different questions. For instance, rather than adopting an Optimum theory across the board, one might endorse an Optimum stance towards the problem of partial acceptance and a Maximizing Expectation stance towards the problem of future similarity. More generally, it is not plausible that we should adopt the same stance to every question. For example, a Fixed approach has *prima facie* plausibility in the face of partial acceptance concerns, but perhaps less plausibility in the face of future similarity concerns. Why should we focus only on futures that are not exactly similar to ours, but also not very different? This is less compelling than the intuition that, in the face of partial acceptance concerns, we should consider how codes perform at around the ninety percent rate.

The one view that does seem plausible across the board is a Maximizing Expectation style theory. This view has many benefits in the partial acceptance context, and also seems plausible when considering questions from methods to future similarity. In the face of uncertainty, the Rule-Consequentialist should simply perform an expected value calculation that takes into account the relative probabilities of the uncertain parameter obtaining.

IV. IS RULE-CONSEQUENTIALISM IRREPARABLY ARBITRARY?

The previous argument suggests a commonality among debates about Rule-Consequentialism. This indicates new ways in which to define the Ideal Rules, with varying assumptions about various topics. However, the flip side to this commonality is that it also reveals the breadth of Rule-Consequentialism's potential problems. Does the range of Rule-Consequentialism's required assumptions imply a death by arbitrariness? If there is no principled way to choose among the resulting expanded list of definitions for the Ideal Rules, Rule-Consequentialism cannot provide a principled defense of any set of moral rules.[22]

One way to explain the objection is to note that expanding the possible topics about what we need to make assumptions about for defining the Ideal Code expands the possibility for arbitrary assumptions. Does Rule-Consequentialism's initially plausibly idea, that morality is to be understood as a welfare maximizing set of rules, collapse under the weight of arbitrariness?

Although acknowledging the broader problem of Rule-Consequentialism's assumptions increases the required number of Rule-Consequentialism's theoretical commitments, we do not yet need to conclude that the theory collapses under the weight of arbitrariness. For one, it is not clear that a theory's increasing detail or sophistication—or decreasing simplicity—should count against it. Insofar as theorists offer persuasive reasons and arguments for Rule-Consequentialism's assumptions, the theory does not collapse under arbitrariness.

Second, the problem of arbitrariness might be dismissed by appeal to reflective equilibrium[23] or some other methodology. Although refining Rule-Consequentialism might seem to involve arbitrary decisions, we can conduct that process through a robust philosophical method. In so far as the inquiry is conducted by reflective equilibrium, or another methodology, this provides another reason against concerns of arbitrariness.

More broadly, the arguments about how to refine Rule-Consequentialism do not involve arbitrary reasons. All plausible moral theories face choices. Take Consequentialism more broadly. Should we endorse act or rule consequentialism;[24] actual or expected consequentialism;[25] scalar or non-scalar versions;[26] or other versions?[27] These are debates that advance reasons and arguments, not merely arbitrary preferences.

---

[22] Thanks to an anonymous referee for suggesting this question.

[23] See, for example, Hooker, *Ideal Code, Real World.*

[24] Richard Brandt, *Morality, Utilitarianism, and Rights.*

[25] Marcus G. Singer, 'Actual Consequence Utilitarianism', *Mind* 86 (1977), pp. 67–77; Peter Railton, 'Alienation, Consequentialism, and the Demands of Morality', *Philosophy and Public Affairs* 13 (1984), pp. 134-71.

[26] Alistair Norcross, 'Comparing Harms: Headaches and Human Lives', *Philosophy and Public Affairs*, 26 (1997), pp. 135–67; Kevin P. Tobia, 'A Defense of Scalar Utilitarianism', *American Philosophical Quarterly* 54 (2017), pp. 283-293.

[27] Not to mention aggregative consequentialism (R. Chao, 'Aggregative Consequentialism', *Southwest Philosophical Review* 31 (2015), pp. 125-136); biocentric consequentialism (R. Attfield, 'Biocentric Consequentialism and Value Pluralism: A Response to Alan Carter', *Utilitas* 17 (2005), pp. 85-92); character consequentialism (P.J. Ivanhoe, 'Character Consequentialism: An Early Confucian Contribution to Contemporary Ethical Theory', *Journal of Religious Ethics* 19 (1991), pp. 55-70); combinative consequentialism (J.E. Gustafsson, 'Combinative Consequentialism and the Problem of Act Versions', *Philosophical Studies* 167 (2014), pp. 585-596); commonsense consequentialism (D.W. Portmore, *Commonsense Consequentialism: Wherein Morality Meets Rationality* (2011)); global consequentialism (P. Pettit and M. Smith, 'Global Consequentialism', in *Morality, Rules and Consequences: A Critical Reader* (B. Hooker, E. Mason, and D. Miller eds., 2000); justicized consequentialism (S. Wigley, 'Justicized Consequentialism: Prioritizing the Right or the Good?' *Journal of Value Inquiry* 46 (2012), pp. 467-479); Kantian consequentialism (D. Cummiskey, 'Kantian Consequentialism', *Ethics* 100 (1990, pp. 586-615); minimal consequentialism (P. Caws, 'Minimal Consequentialism', *Philosophy* 70 (1995), pp. 313-339); multi-dimensional consequentialism (M. Peterson, *The Dimensions of Consequentialism* (2013)); multiple-act consequentialism (J. Mendola, J., *Goodness and Justice: A Consequentialist Moral Theory* (2006)); person-based consequentialism (M.A. Roberts, 'Can it ever be better never to have existed at all? Person-based consequentialism and a new repugnant conclusion', *Journal of Applied Philosophy* 20 (2003), pp. 159-185); progressive consequentialism (Jamieson, D. & Elliot, R, 'Progressive Consequentialism', *Philosophical Perspectives* 23 (2009), pp.: 245-251); and virtue consequentialism (B. Bradley, 'Virtue Consequentialism', *Utilitas* 17 (2005),

In light of all these considerations, I am most inclined to retain cautious optimism about Rule-Consequentialism's project. The mere existence of these questions—how should Rule-Consequentialism respond to various uncertainties—does necessarily indicate Rule-Consequentialism's arbitrariness, nevermind its failure. Of course, it might turn out that with respect to some of these questions there is no clear victor among the various versions of Rule-Consequentialism. That conclusion is far from settled, and given the numerous principled arguments in favor of these various approaches (e.g. in the partial acceptance debates), for now the burden of proof seems to rest with those who think Rule-Consequentialism is overburdened by arbitrariness.

V. CONCLUSION

The discussion about Rule-Consequentialism's "problem of partial acceptance" initially appears narrow. It begins with a simple question—how should the ideal code be selected given the possibility that its rules may not be universally accepted?—and results in an array of complex theoretical responses: Fixed Rate, Top Rate, Variable Rate, Optimum Rate, and Maximizing Expectation Rate Rule-Consequentialism.

That problem has no easy solution. The failure of "Calculated Rates" Rule-Consequentialism demonstrates this difficulty. But it also illuminates the broader problem underlying the problem of partial acceptance. Formulating the best version of Rule-Consequentialism requires grappling with the theory's assumptions about various uncertainties.

This new problem—how should the ideal code be selected given the possibility that various assumptions will not be (fully) realized?—is broad. It implicates concerns about partial compliance, partial acceptance, ineffective training and enforcement, the dis/similarity of the future, and perhaps further considerations. Though seemingly narrow, the previous rich discussions about partial acceptance provide a taxonomy and groundwork for broader attempts to formulate the best version of Rule-Consequentialism.

This is not to say that these approaches solve all of Rule-Consequentialism's problems. Future work could delve into different details. For example, even in the partial acceptance debate, Fixed theory might have very different implications depending on *which* ninety percent of the population accepts the rules. For at least some rules, a dense pocket of dissent plausibly has very different consequences than evenly distributed non-acceptance. Debates among these various theories (e.g. Fixed vs. Optimum vs. Maximizing Expectation) should come alongside debates within each theory.

Kevin.Tobia@yale.edu

---

pp. 282-298).