

How to Be a Rule-Utilitarian: Introducing Variable-Rate Rule-Utilitarianism¹

Rule-utilitarianism, in spite of its considerable attractions, is a theory in need of a plausible and precise formulation. The basic idea behind rule-utilitarianism is that right action should be defined in terms of what would be required by rules which would maximise either actual or expected utility if those rules gained general acceptance or (on some versions of the theory) general compliance. Rule-utilitarians differ over whether acceptance or compliance is the key notion (see Hooker 2000: 75-80) and also over whether the theory should be couched in terms of actual or expected utility. I shall here officially remain neutral on these difficult questions, as the main arguments here go through equally well on either interpretation, assuming that both versions of the theory are otherwise defensible, anyway (see below). However, partly for ease of exposition and partly because I consider it to be the more plausible view, I shall frame my discussion in terms of acceptance and expected utility.

One advantage of the acceptance-based version of rule-utilitarianism worth noting, though, is that it makes it easier to block the classic objection that rule-utilitarianism collapses into act-utilitarianism. One version of this objection maintains that there is precisely one rule which is such that universal compliance with it would maximize utility, namely the principle of utility itself. Even this is not obvious. In a famous discussion, Donald Regan calls our attention to cases in which universal compliance with the principle of utility need not maximize utility when compared with an alternative and available set of actions (see Regan 1980). The point there is that there can be cases in which the consequences of each of our actions depends on what the other will

in fact do. He gives the example of Whiff and Poof who are causally insulated from one another (and so cannot influence each other's decisions) and who each have the simple choice of pressing a button or not. If they both press their respective buttons then ten units of utility result, if neither press then six units of utility result, while if only one presses his button then zero units of utility are produced. It seems that if both press their buttons then they have both complied but if neither press their button then they have both complied too (given that they must each take the action of the other as fixed). This suggests that universal compliance with the principle of utility need not maximize utility (though it can; they could have both pressed their buttons). The issues raised about acting as part of a group are complex and fascinating, but I must here put them to one side. The main point is that even a version of rule-utilitarianism couched in terms of general compliance may have more resources than one might have imagined to block the collapse objection.

In any event, couching the theory in terms of acceptance rather than compliance provides considerably more and richer resources with which to block the collapse objection. For acceptance of the principle of utility even by those who are conscientious need not lead to compliance with the principle. Even those who make a good faith effort to comply with the principle may get the relevant facts wrong and even the best of us can sometimes be prone to special pleading or self-deception about our moral duties. Also, the general acceptance of the principle of utility as a guide to day to day decision-making would eventually be recognized as such and this in turn would reduce people's willingness to trust one another. For how can I depend on you to keep your promise

¹ Many thanks to Brad Hooker, Elinor Mason and the participants of the staff seminar at Leeds University for useful discussion of an earlier draft of this material.

when I know you will break it if the sums (as you compute them) come out one way rather than another? Finally, we must bear in mind the costs of internalizing the principle of utility as a guide to daily decision-making, since getting someone to internalize the principle involves not just getting them to memorize it (that would be easy!) but also getting the principle to play a role in their motivational economy. With a principle as demanding as the principle of utility, this might be very costly indeed, given the natural partiality of human beings. For an elaboration of just these points and a more in-depth discussion of the collapse objection, see Hooker 2000: 92-99. For these reasons, I do think the acceptance-based version of rule-utilitarianism is more promising.²

The rough idea behind utilitarianism has a lot to recommend it, but so far it is too rough to assess. The difficulty I explore here concerns how we should understand 'general acceptance' (or 'general compliance'). Just how much acceptance is necessary for general acceptance? The question is not of merely theoretical interest, since different answers may well generate a different set of moral duties in the real world. Since one of the attractions of utilitarianism is its apparent ability to help us make difficult decisions on a principled basis it would not be satisfactory to leave such a key concept irreducibly vague and open-ended. Therefore, defenders of rule-utilitarianism have typically tried to make the idea more precise. Here is where the real trouble begins.

It seems that rule-utilitarians face a dilemma. Either they characterize general acceptance as 100% acceptance or characterize it as something less than 100% acceptance. On the first horn of the dilemma the rule-utilitarian is vulnerable to the charge of utopianism while on the second horn of the dilemma she is open to the charge of arbitrariness and a lack of philosophical depth. The problem with characterizing the

² Thanks to an anonymous referee for drawing me out more on this point.

theory in terms of 100% acceptance of the rules is that in the scenario of 100% acceptance there will be no need to incorporate rules to deal with problems uniquely generated by non-acceptance.

As Brad Hooker and others have pointed out, non-acceptance of the moral rules generates its own moral problems. Contrast the criminal who accepts the moral rules but nonetheless gives in to temptation and commits a crime. The best strategy for rehabilitating this sort of criminal may be very different indeed from the best strategy for dealing with a criminal who does not accept the moral rules in the first place- either a socio-path with no moral conscience or someone who accepts deviant moral norms which contradict the ideal code (Hooker 2000: 82). Similar points apply in the case of deterrence. So the non-acceptance of the ideal moral rules does generate its own unique problems above and beyond mere non-compliance. However, these costs will not arise in a world of 100% acceptance.

Moreover, there will be some costs associated with incorporating such further rules to deal with problems associated with non-acceptance – teaching children more rules, learning them, and so on have various social costs. So a moral code which is ideal for a world of 100% acceptance of that code would have no rules for dealing with non-acceptance, for such rules would be pointless and at least somewhat costly. However, in the real world the phenomenon of those who do not accept an ideal moral code is all too real and generates problems which call for action. So a moral theory which provided no guidance for such situations would to that extent be implausibly utopian.

Most rule-utilitarians have avoided this horn of the dilemma by specifying a particular level of social acceptance which is less than 100% and formulating their theory

in terms of this level. Richard Brandt and Brad Hooker both define right action in terms of acceptance by 90% of the relevant population (see Brandt 1992 and Hooker 2000). Defenders of rule-utilitarianism do seem slightly embarrassed by this feature of their theory since they seem to be pulling these numbers out of thin air, and there does seem to be something slightly odd about the very idea that the fundamental principle of morality might intuitively be arbitrary in this way. However, the idea seems to be that we should learn to live with a little arbitrariness at the margin if the resulting theory fits very well with our considered moral judgements in a moment of cool reflection. If the only problem were arbitrariness as such then this might be a reasonable albeit still slightly embarrassing (can the ultimate principle of morality *really* be intuitively arbitrary in this way?) defence. However, the problem goes deeper than this for at least two reasons.

First, fixing a specific level of social acceptance which is less than 100% actually faces a utopianism objection very similar to the one facing full acceptance versions of the theory. The problem with theories framed in terms of 100% acceptance is that they were simply not designed to deal with situations in which there is less than 100% acceptance of the ideal rules and, after all, in the real world we may well find ourselves in such circumstances. An isomorphic objection can be pressed against those rule-utilitarians who formulate their theory in terms of some specific level of social acceptance which falls short of 100%. For a theory framed in terms of (e.g.) 90% acceptance is simply not designed to deal with situations in which there is less than 90% acceptance of the ideal rules and, after all, in the real world we may find ourselves in just such circumstances.

Consider again the standard charge of utopianism against theories couched in terms of 100% acceptance. Such theories will generate no rules to deal with problems

generated by those who do not accept the ideal rules, such as the need for different strategies of deterrence and rehabilitation, e.g. Now consider a theory framed in terms of 90% acceptance. Such a theory will generate some rules to deal with non-acceptance of the ideal code and this is an improvement over full-compliance versions of the theory. However, the rules generated by such a theory will be narrowly tailored to deal with the problems generated by non-acceptance of the rules *by exactly 10% of the population*. We may need very different rules to cope with situations in which non-acceptance of the theory is very much more widespread than this.

Suppose that only 60% of the population accepts the ideal code. In that case, the problems generated by non-acceptance will be much more widespread and hence much more serious and urgent and we may well need very different rules from those that would be adopted in the 90% acceptance scenario to deal with these problems. For example, in a world of 90% acceptance a rule requiring people to engage in a certain amount of moral proselytising may simply be a recipe for costly (in utilitarian terms) clashes between people with very different values. In such circumstances, the utilitarian case against busybodies and moral arrogance should not be that difficult to make out. However, what in a world of 90% acceptance might be moral arrogance might well be a necessary evil in a world of 60% acceptance, depending of course on the consequences of such proselytising. In such dire circumstances, the costs associated with unpleasant confrontation and argument may be outweighed by the benefits of increasing acceptance of (or at least compliance with) the moral rules.

Of course, it is ultimately an empirical question whether moral proselytising would be likely to work anyway, but that is enough to make the point. Intuitively it

should be an empirical question which hinges on the effectiveness of proselytising and the costs associated with such interference. If we let the rules designed for a world of 90% acceptance be our guide in these more dire circumstances, though, such moral reasons will become invisible regardless of how these empirical facts turn out. Once again, this seems utopian. Why should the fact that we clearly would not need rules requiring (or perhaps even allowing) extensive moral proselytising determine whether such proselytising is appropriate in less happy circumstances? So although a version of rule-utilitarianism which specifies a level of acceptance less than 100% can reduce the force of the original charge of utopianism, the charge still stands.

The second problem with specifying a particular level of social acceptance less than 100% is that such a theory lacks explanatory depth. Rule-utilitarianism traditionally aspires to provide the ultimate principle of morality; it should be an (the) axiom and not a theorem. However, there are very good reasons to think that any version of the theory couched in terms of a specific level of acceptance (less than 100%) will at best be a theorem and not an axiom, thus failing to provide the level of explanatory depth to which rule-utilitarians rightly aspire. To see why this is so, just pause to ask yourself why the figure of 90% seems to be roughly in the right ballpark (assuming, that is, that you do think it is roughly correct; otherwise substitute what you take to be the most plausible candidate alternative level of acceptance). It seems clear on reflection that 90% seems plausible to us insofar as it does because 90% seems like a realistic ideal *for us*. In other words, it does not seem implausible or counter-intuitive to suppose that human beings might eventually be able to arrange their affairs so that roughly 90% of the population accepted the ideal moral code so long as that code is not incredibly demanding. By

contrast, 99% acceptance seems like a very unrealistic ideal for creatures like us and suitable instead for angels or Vulcans (like Mr. Spock from *Star Trek*). At the other extreme, levels of social acceptance much lower than 90% are realistic enough for creatures like us, all right, but is not sufficiently ambitious. We can do better than *that*, one is inclined to think. The crucial point is that the tenability of these reactions depends in large part on implicit assumptions about human nature. A species of creatures perfectly capable of generating and sustaining very high levels of acceptance might (like Vulcans or angels) well find 99% acceptance a plausible ideal and consider 90% acceptance insufficiently ambitious. Whereas a species of creatures much more prone to selfishness or violence (a species of congenital scoundrels) might rightly consider 90% acceptance to be wildly optimistic and utopian.

The upshot of all this is that what level of social acceptance is germane for a given group of people will vary from one group to the other in a systematic way depending on the psychologies of the members of the groups in question. There ought to be some principled explanation of this striking co-variation. However, if there is a deeper moral principle which provides a function from facts about the psychologies of a group of creatures (and perhaps other related facts) to a level of social acceptance which is suitable for a moral principle governing such creatures then a rule-utilitarian theory couched in terms of the level of acceptance which just happens to be suitable for human beings (90% e.g.) will turn out to be a mere theorem. The ultimate moral principle will not be this locally correct (we are now assuming) principle but instead a deeper principle which provides a function from psychological facts about the group under consideration and

perhaps other facts to a version of rule-utilitarianism specified in terms of a particular level of social acceptance.

One reply to this objection holds that, “rule-utilitarianism is a relativistic doctrine” which “seeks the ideal rules for a particular society...if we can count on a 90% compliance [acceptance] rate in our society, then, given that we are developing rules for our society, 90% is not an arbitrary figure.” However, not all rule-utilitarians understand their theory in relativistic terms. Brad Hooker, for example, casts his theory in terms of acceptance by the majority of “everyone everywhere, rather than just everyone in the agent’s society.” (Hooker 2000: 83). However, other rule-utilitarians might be happy with a relativistic account. A further and deeper problem with this line of reply is that what level of acceptance of the rules we can realistically accept depends on the content of the rules. However, in order to fix the content of the rules we must already have some idea of the level of acceptance. For the content of the rules is determined by which rules would maximize utility *if accepted by such-and-such percentage of the population*. So it seems that we must know which rules would be ideal to know what the right level of acceptance is for a given society, but we must know what the right (realistic) level of acceptance for a given society is before we can know which rules count as ideal in the relevant sense. It is not clear how this sort of vicious circularity can be avoided in the present framework, which is part of what motivates the alternative developed in the text.³

It would be interesting to see how a rule-utilitarian might try to avoid this worry by developing a theory which does include a principled function from psychological facts about a given population (and perhaps other facts) to a rule-utilitarian theorem couched in terms of a level of social acceptance suitable to agents of whom such facts hold. My

hunch is that this project would be very difficult to carry out with anything like the precision to which utilitarians have traditionally aspired and without running into further difficulties of a rather different sort, but since this is just a hunch I shall leave it at that. However, there is another very different solution available to rule-utilitarians which departs more radically from the standard formulations in the literature and seems more promising.

The problems for rule-utilitarianism discussed here derive from the rule-utilitarian's apparent need to specify a particular level of social acceptance for the ideal code. Perhaps we should re-examine the assumption that rule-utilitarians really are committed to providing a specific level of acceptance. A plausible form of rule-utilitarianism does need to avoid the charge of utopianism, but the simplest solution to this problem is not to move from a theory couched in terms of 100% acceptance to a theory couched in terms of some specific level of acceptance which is less than 100%. For we can instead reject the more basic idea that rule-utilitarianism needs to be formulated in terms of any specific level of social acceptance without making our theory hopelessly vague and indeterminate. Perhaps, in other words, we should reject what we might call the 'fixed-rate' interpretation of rule-utilitarianism which insists on privileging some specific level of acceptance for a given society. Instead of privileging one specific level of social acceptance we could in effect include all possible levels of social acceptance in our account of right action. In particular, we could hold that an action is right just in case it would be required by rules which have the following property: when you take the expected utility of every level of social acceptance between (and including) 0% and 100% for those rules and compute the *average* expected utility for all of those

³ Thanks to an anonymous referee for pressing me on this point.

different levels of acceptance, the average for these rules is at least as high as the corresponding average for any alternative set of rules. Call this account of right action “variable-rate rule-utilitarianism.” The variable-rate approach has a number of important advantages over more traditional “fixed rate” approaches.

First, variable-rate rule-utilitarianism does not seem arbitrary in the way that theories which specify a particular level of compliance less than 100% (e.g. 90%) do. Second, unlike the utopian 100% approach, variable-rate rule-utilitarianism arguably will generate rules or social practices (see below for a more sophisticated implementation of the basic idea which goes via a division of labour) to cover problems arising from non-acceptance. For a set of rules which included no such rules would do well only in the scenario of 100% acceptance and would do very badly compared to other sets of rules in circumstances of less than 100% acceptance. Therefore such a set of rules would be very unlikely to have the highest average expected utility across all levels of social acceptance. Third, unlike approaches which fix a specific level of social acceptance which is less than 100% (e.g., 90%) the proposed account will not generate a set of rules which includes no useful prescriptions for levels of acceptance which fall well below 90% (or whatever).

While it may not be immediately obvious, there is a good prima facie argument that the proposed account would generate a set of rules which are conditional in form and which make different demands depending on how widely accepted the moral code is. So the ideal code might include rules of the form, “If roughly 90% of the population accept this code then do not engage in moral proselytising of such-and-such form” but also include rules of the form, “If roughly 60% of the population accept the code then engage in such-and-such sort of proselytising.” Such a set of rules should have a higher average

across different levels of social acceptance than a rival set of rules which is pegged to just one level of social acceptance simply in virtue of its greater flexibility. Of course, there will be some costs associated with the greater complexity of the rules. They will be harder to teach and pass along to the next generation, e.g. and also a good deal harder to apply (how does one determine exactly how widely accepted a given moral code is in one's society?). However, these costs could easily be outweighed by the benefits of having rules which would help people deal much more effectively with a much wider range of circumstances. Moreover, this reflects the pre-theoretically attractive idea that an ideal moral code should be flexible enough to deal with a wide range of circumstances, where the variation in circumstance can include variation in the extent to which the code itself is accepted. Intuitively, an ideal moral code is one which would be helpful in a wide range of circumstances.

Perhaps the foregoing reply to the concern about greater complexity of the rules generated by my proposal is too swift. Here it is worth being clear about the overall dialectic. Suppose that the internalization costs of conditional rules keyed to each possible level of acceptance are actually very high. This may either mean that the ideal code will have no such conditional rules or, more likely, that it will have a fairly limited set of them which is not nearly as fine-grained as is logically possible but which still retains a great deal of the virtues of flexibility discussed above. In the former case (if the ideal code includes no conditional rules) then the argument from conditional rules in favour of my proposal must be conceded to be unsound but the remaining arguments for the proposal remain intact. In the latter case (if the ideal code contains some conditional rules but far fewer than would be needed for being as fine-

grained as could matter to the utility of complying with the rules) then we can retain both the position and the argument from conditional rules, though this argument will now have to be a bit more complex. For example, perhaps the ideal code would include rules for conditions in which 'virtually everyone accepts the rules', where 'most people accept the rules', in which 'at least a majority accept the rules' and where 'less than half the people accept the rules' and where 'virtually nobody accept the rules'.

Ultimately, of course, it will be an empirical question just which rules would be generated for a given population by the proposed account. In any event, this compromise position would go some way to meeting concerns about costs of internalization of conditional rules without abandoning the argument from conditional rules in favour of the approach. Moreover, this approach also strikes me as more psychologically realistic and easy to follow than rules keyed to much more precise levels of social acceptance. Aristotle's remarks on not importing more precision to the discipline that it sensibly implies is perhaps germane here. However this pans out, the main dialectical point to bear in mind is that the question raised by the costs of internalising conditional rules is an objection only to one argument for my proposal and not an objection to the proposal itself nor to the other arguments given for that proposal.

A slightly different objection, which itself suggests a more promising way of developing the theory, goes as follows. Imagine a society that accepts a set of rules R with an acceptance level of 90% and which is ideal given 90% acceptance. On the theory on offer, this may well not be enough. For such a society apparently will also need to accept some conditional rules to cover other possible levels of acceptance. Perhaps these rules will be very complex, in which case accepting these rules might force the members

of this population to accept an oversimplified version of R just in order to be able to handle the complexity of all the conditional rules. There could be substantial costs associated with such oversimplifications of R, in which case it seems like the members of the population will have to learn a set of rules that, given their actual acceptance rate, is inferior, and this seems strange.

One reply to this objection is that it seems to assume that the acceptance rate for the population in question is fixed. Otherwise, the conditional rules for dealing with backsliding (and moral progress) might have considerable expected utility. Since no realistic human society (much less everyone, everywhere, which is the scope of Hooker's own formulation) is likely ever to have a fixed rate of acceptance of the ideal moral rules, the objection might require giving great weight to fairly strange circumstances. The correct rejoinder to this reply is that even if there is substantial utility associated with the conditional rules (those rules conditional on levels of acceptance, that is), this may not be enough to outweigh the utility lost associated with an oversimplification of their non-conditional rules (non-conditional on levels of acceptance, that is).

A second, and better, reply to the original objection is simply that if the conditional rules are sufficiently coarse-grained then it will be very unlikely that internalizing them would force the population to accept an oversimplified version of the non-conditional rules. Very coarse grained rules are, after all, not that difficult to internalize. The issue then becomes whether such-coarse-grained rules are likely to be the ones which would maximize average utility across different acceptance rates. Insofar as the original objection has any force, there will be some substantial costs associated with more fine-grained rules. In addition to being costly to internalize and difficult to

remember, they will also (according to the objection) force people to oversimplify their non-conditional moral rules in ways that have significant costs. This in itself suggests that it would be better across most levels of acceptance to have people embrace more coarse-grained conditional rules than more fine-grained ones.

A third, and even better reply, is to emphasize that morality should be understood as a collective affair. As such it can and should involve a division of labour. Indeed, this idea is not so alien to ordinary practice. When deliberating on difficult and novel issues in bioethics (e.g.) we do often rely on the testimony of designated experts in that field. The suggestion here is that ideally we would treat large-scale shifts in acceptance of the moral rules as a problem to be at least partially handed over to experts as well. So perhaps the ideal rules that we should all accept would counsel all of us to rely on the judgments of a subpopulation (possibly quite small given the overall population) experts as to how we should revise our moral code if levels of acceptance suddenly shift in some significant way. Insofar as we all accepted such a moral code, we would be motivated to provide the infrastructure needed for such a division of labour and could therefore reap its benefits. One such benefit which is highly germane here is that the great mass of humanity would not be required to internalize any rules at all about how we should revise our moral code if there was a big shift in acceptance levels. For we could instead rely on the advice of those experts who we have nominated to deliberate long and hard about such matters. Of course, such reliance would need to be defeasible; if the 'moral experts' counselled something that seemed morally outrageous then the rest of us would be warranted in being very cautious indeed about implementing their recommendations.

Moreover, the proposal should not be understood as anti-democratic; the testimony of the experts should be understood on the model of advice that we should take very seriously but in no way legally binding on the rest of us. This is not government house utilitarianism even about the fairly limited question of how we should revise our moral code in light of shifting levels of acceptance. If such a scheme is workable then it seems likely to have the virtue of flexibility in the face of shifting levels of acceptance without the down side of being forced to oversimplify the non-conditional elements of its moral code to achieve this level of flexibility.⁴

A further objection begins with the premise that in any population larger than one person, most levels of social acceptance (all of those other than 0% and 100%) will be multiply realizable in terms of different distributions of acceptance. For example, in a population of two people, call them A and B, 50% acceptance will be realizable in two ways: (1) A accepts and B does not and (2) B accepts and A does not. This suggests a variant on the account proposed in the text formulated not merely in terms of all levels of acceptance but instead couched in terms of the different realizations of different levels of acceptance. Why, one might wonder, should we prefer the account sketched in the text to this alternative? Moreover, if there is no principled grounds for choosing between these alternatives then is not my own account vulnerable to an arbitrariness objection just as much as Hooker's 90% based account? For just as Hooker arbitrarily 'plumps for' 90% social acceptance, I 'plump for' levels of acceptance instead of all of the different realizations of those levels of acceptance.

⁴ Many thanks to an anonymous referee for pressing me on this point, which has indeed enriched my own understanding of the problems for and resources of the account on offer.

Even if this objection were sound, it would undermine only one of the main arguments for the proposal – the argument from arbitrariness. The argument from utopianism would remain intact. However, it seems to me that the objection is not sound. For it is not arbitrariness but good sense for one’s moral theory to be no more fine grained than the subject matter requires. Different levels of social acceptance matter because it is so plausible that any realistic moral community will need moral rules to deal with non-acceptance of the ideal code. By contrast, there is no realistic need for a set of moral rules which is sensitive to particular realizations of different levels of acceptance. Indeed, on some ways of understanding morality it is a priori that moral rules (as opposed to legal rules or rules of etiquette) do not include proper names, which rules couched in terms of particular realizations of each level of acceptance presumably would. So there is a principled rationale for moral theory being sensitive to different levels of acceptance which does not carry over to different realizations of those levels. Hence the distinction drawn by the proposed theory is not arbitrary. In which case it is not arbitrary to prefer the theory outlined in the text to one couched in terms of different realizations of different levels of acceptance.⁵

As far as I can see, variable-rate rule-utilitarianism has also has all of the advantages associated with fixed-rate versions of rule-utilitarianism but without the costs discussed above. So there is a dominance argument for variable-rate rule-utilitarianism over its more traditional fixed-rate rivals. Of course, I do not pretend to have shown that this is the only way to avoid the problems arising from the assumption that a particular level of social acceptance should be privileged by the rule-utilitarian. There may well be other interesting ways to avoid the problems that I have not even considered. Nor do I

⁵ Many thanks to Peter Millican for pressing me on this point.

mean to suggest that in this short article I have solved all of the problems facing rule-utilitarianism. My contention is merely that these remaining problems shall be problems for fixed-rate versions of rule-utilitarianism (like Hooker's and Brandt's) at least as much as they are for variable-rate versions. However, unless some superior solution to the problems of arbitrariness and utopianism canvassed here can be found we should formulate rule-utilitarianism in terms of average utilities of rules across all of the different levels of social acceptance of those rules. The resulting theory is perhaps more complex than traditional formulations of rule-utilitarianism, but in this case the extra complexity buys us a great deal of plausibility in terms of avoiding otherwise damning objections.

If the proposal discussed here is defensible then rule-utilitarians can avoid the vagueness of 'general acceptance' left unexplicated, the utopianism of couching the theory in terms of 100% social acceptance, and the arbitrariness and associated objections facing any version of the theory which specifies a level of social acceptance less than 100%. For those wanting a clear statement of the positive (dominance) argument for variable-rate rule utilitarianism, it is this: Variable-rate rule-utilitarianism has all the advantages of fixed-rate rule-utilitarianism (for a survey of these, see Hooker 2000) while avoiding some rather serious disadvantages of the fixed-rate approach. So insofar as the standard advantages of rule-utilitarianism are significant, we should embrace variable-rate rule-utilitarianism on their strength rather than its fixed-rate rival.

Bibliography

Brandt, R. 1992. *Morality, Utilitarianism, and Rights*. Cambridge: Cambridge University Press.

Hooker, B. 2000. *Ideal Code, Real World*. Oxford: Oxford University Press.

Regan, D. 1980. *Utilitarianism and Co-operation*. New York: Oxford University Press.